On Stable and Efficient Mechanisms for Priority-based Allocation Problems*

Kang Rong, Qianfeng Tang,† Yongchao Zhang

School of Economics, Shanghai University of Finance and Economics, Shanghai, 200433, China

January 6, 2020

Abstract

For school choice (priority-based allocation) problems, when the priority structure is acyclic, the associated student-proposing deferred acceptance algorithm is Pareto efficient and group strategy-proof (Ergin, 2002). We reveal a hidden iterative removal structure behind such deferred acceptance algorithms. A nonempty set of students is called a *top fair set* (TFS) if when all students apply to their most preferred schools and all schools accept the best applicants up to their quotas, students in the set are always accepted, regardless of other students' preferences. We provide an elimination process to find the maximal TFS, if any TFS exists. We show that for any priority structure,

*We are grateful to the Editor, two anonymous referees, Yajing Chen, Yi-Chun Chen, Xiang Han, Audrey Hu, Zhiwei Liu, Mengling Li, Shravan Lucraz, Ning Sun, Utku Ünver, Siyang Xiong, Jingyi Xue, Zhixiang Zhang, seminar participants at Capital University of Economics and Business, Central University of Finance and Economics, City University of Hong Kong, East China University of Science and Technology, Peking University, Singapore Management University, University of Nottingham Ningbo, Xiamen University, and participants at the 4th Microeconomics Workshop in Nanjing, the 9th Shanghai Microeconomics Workshop, Chengdu Economic Theory Workshop (SWUFE), Workshop on Matching, Search and Market Design at National University of Singapore (IMS) for helpful comments and discussions. This project is supported by National Natural Science Foundation of China (No. 71873086, No. 71803114, and No. 71873081) and Program for Innovative Research Team of Shanghai University of Finance and Economics (No. 2017110093).

†Corresponding author.

Email addresses: rong.kang@mail.shufe.edu.cn (Rong), tang.qianfeng@mail.shufe.edu.cn (Tang), yongchao@mail.shufe.edu.cn (Zhang)

iterative removal of TFS is equivalent to the associated deferred acceptance algorithm if and only if the latter is a Pareto efficient mechanism.

JEL classification: C78; D61; D78; I20

Keywords: Deferred acceptance algorithm, Pareto efficiency, school choice, stability, strategy-proofness

1 Introduction

We study the allocation of a set of indivisible objects to a set of agents without monetary transfers. Each agent has unit demand and is assumed to have strict preference over object types. The most prominent example of such problems is school choice.[1] We will refer to agents as students, object types as schools and study school choice problems. An allocation mechanism specifies an assignment of school seats to students for each profile of students' preferences. In practice, each school has a fixed quota of seats and is often associated with a priority list over students. Schools' priority lists and quotas together define a priority structure.

Stability and students' welfare are both desirable in school choice. The concept of stability is introduced in the classical work of Gale and Shapley (1962) for college admission problems, and it is later reinterpreted as fairness in school choice (Balinski and Sönmez, 1999; Abdulkadiroğlu and Sönmez, 2003). An assignment is fair if it eliminates priority violation (or justified-envy), i.e., if there is no student-school pair in which the student desires the school while a lower-priority student is assigned. However, it is well known that fairness and Pareto efficiency are not compatible: even the student-optimal stable matching, which is produced by Gale and Shapley's celebrated student-proposing

[1] Other examples include house allocation, task assignment, and course assignment.

deferred acceptance algorithm (henceforth, DA), may not be Pareto efficient for students (Roth, 1982; Abdulkadiroğlu and Sönmez, 2003). Ergin (2002) characterizes priority structures at which DA is an efficient mechanism and calls them acyclic priority structures.[2]

We study the driving force behind efficient DA mechanisms by revealing an iterative structure behind their deferred acceptance appearance. We call a nonempty set of students a *top fair set* (TFS) if when all students apply to their most preferred schools and all schools accept the best applicants up to their quotas, students in this set are always accepted, regardless of other students' preferences. For a given school choice problem, a TFS may or may not exist, and when it exists, we can assign students in it and remove them. After that, iteratively, we search for TFS at the remaining subproblem, and upon existence, assign and remove it. We call this process the TFS algorithm.

A priority structure is said to be *TFS-solvable* if for any profile of students' preferences, a TFS always exists in each step of the TFS algorithm, until all are assigned. Our main result shows that, a priority structure is TFS-solvable if and only if the associated DA mechanism is Pareto efficient, and when that happens, the TFS algorithm is equivalent to the DA mechanism. Ergin (2002) shows that for any DA mechanism, the acyclicity of the priority structure, Pareto efficiency, group strategy-proofness, and consistency are all equivalent. Therefore, a priority structure is TFS-solvable if and only if it is acyclic. By decomposing a Pareto efficient DA mechanism into a sequence of TFS, we reveal an iterative removal structure in it and bring more intuition to its properties.

We also discover two properties of TFS. First, the union of any pair of TFS is also a TFS. As a result, there is a maximal TFS if any exists. We then design an iterative elimination process which always find the maximal TFS, if any exists. Second, TFS embeds a

[2]Ergin's acyclicity condition is extensively studied in the literature; e.g., it has been used in the characterization of efficient priority rules (Ehlers and Klaus, 2006), Nash implementation of the stable matching correspondence (Haeringer and Klijn, 2009), and robust stability (Kojima, 2011).

form of consistency in it, just like that of the efficient DA mechanisms. If we assign and remove any subset of students in a TFS, then at the subproblem that remains, the rest of the TFS remains a TFS. With these properties, we know that in any step of the TFS algorithm, when TFS exists, we are free to pick either the maximal one or any part of it to assign.

Beside efficient DA mechanisms, all known efficient and group strategy-proof mechanisms are based on Gale's top trading cycles (henceforth, TTC).[3] The discovery of TFS reveals that a new component, which differs from TTC, can also be used to construct efficient and group strategy-proof mechanisms. Since the formations of TFS and TTC are both independent of other students' preferences, in the terminology of Pycia and Ünver (2017), they are both decisive groups.

This paper most closely relates to Ergin (2002). Ergin's acyclicity condition and equivalence results are later generalized by Kumano (2009) and Kojima and Manea (2010), respectively, to acceptant substitutable priority.[4] Kesten (2006) characterizes priority structures at which the priority-based TTC mechanism is stable, which form a strict subset of Ergin's acyclic priority structures. Recently, Abdulkadiroğlu et al. (2019) show that when each school has only one seat, the priority-based TTC mechanism is maximally stable among strategy-proof and Pareto efficient mechanisms. Kesten (2004), Hakimov and Kesten (2018) and Morrill (2015) propose strategy-proof and Pareto efficient variations of the TTC mechanism which may improve its stability. Nonetheless, these variations are not maximally stable, because in general when the priority structure is acyclic, they are

[3]The TTC mechanism is first introduced by Shapley and Scarf (1974) in housing market and is characterized by Ma (1994). TTC-based mechanisms include, in increasing generality, the serial dictatorships (Satterthwaite and Sonnenschein, 1981; Svensson, 1994), the priority-based TTC mechanisms (Abdulkadiroğlu and Sönmez, 2003), the hierarchical exchange rules (Pápai, 2000), and the trading cycles mechanisms (Pycia and Ünver, 2017).

[4]The DA mechanism is exended to substitutable priority by Roth and Sotomayor (1990). Kojima and Manea (2010) characterize the class of all DA mechanisms when schools have acceptant substitutable priority.

not outcome equivalent to DA, or equivalently, because they are not based on TFS.

The rest of the paper is organized as follows. We present notations, basic concepts, and the DA mechanism in Section 2. TFS and its basic properties are introduced in Section 3, and our main result is presented in Section 4. We conclude in Section 5. All proofs are relegated into Appendix A.

2 Preliminaries

2.1 Notations

Fix a set of agents and a set of indivisible object types. For convenience, we represent the set of agents by a set of students $I = \{1, \ldots, n\}$ and the set of object types by a set of schools $S = \{s_1, \ldots, s_m\}$. The quota of $s \in S$, denoted by $q_s \geq 1$, is the number of available seats at s. There is a null school denoted by $\emptyset$, which has unlimited quota. For any finite set X, let $|X|$ denote the number of its elements.

Each school $s \in S$ is associated with a strict priority list $\succ_s$ over students.[5] If student i has higher priority than student j at s, we write $i \succ_s j$; if $I_1, I_2 \subset I$ satisfy that for all $i_1 \in I_1$ and $i_2 \in I_2, i_1 \succ_s i_2$, we write $I_1 \succ_s I_2$. Each student i has a strict preference P_i over schools in $S \cup \{\emptyset\}$ with symmetric extension R_i. If student i prefers (weakly prefers, resp.) school s' to school s, we write $s' P_i s$ ($s' R_i s$, resp.). If $s R_i \emptyset$, then s is said to be acceptable to student i. If s is i's most preferred school, then we say s is her favorite or she **favors** s.

Let $\succ \equiv (\succ_s)_{s \in S}, P = (P_i)_{i \in I}$, and $q \equiv (q_s)_{s \in S}$ denote the priority profile, preference profile, and the vector of quotas, respectively. A **priority structure** consists of a pair $(\succ, q)$,

[5]Therefore, all students are acceptable for each school. Let the priority list at the null school be an arbitrary strict ranking of I.

and a **school choice problem** consists of a triple $(\succ, q; P)$.[6]

An **assignment** μ is a mapping from I to $S \cup \{\emptyset\}$ such that $|\mu^{-1}(s)| \leq q_s, \forall s \in S$, where $\mu(i) = \emptyset$ represents that i is unassigned. Given a preference profile P, assignment ν weakly Pareto dominates assignment μ at P if for all $i \in I, \nu(i) R_i \mu(i)$, and ν Pareto dominates μ if in addition $\nu \neq \mu$. An assignment is **Pareto efficient** at P if it is not Pareto dominated.

Student i **desires** school s at assignment μ if $sP_i\mu(i)$. We say that at assignment μ, student j **violates student** i**'s priority** at school s or (i, s) is a **blocking pair** of μ, if i desires s, $\mu(j) = s$ and $i \succ_s j$. An assignment μ is **fair** if no student's priority at any school is violated at μ, and it is **non-wasteful** if for any school $s \in S \cup \{\emptyset\}$ that is desired by some student at μ, $|\mu^{-1}(s)| = q_s$. An assignment μ is **stable** if it is fair and non-wasteful.

An **allocation mechanism** φ selects an assignment $\varphi(P)$ for all P. Let $\varphi(P)(i)$ denote i's assignment at $\varphi(P)$. An allocation mechanism φ is Pareto efficient if it always selects a Pareto efficient assignment, and φ is stable w.r.t. a priority structure $(\succ, q)$ if for each $P, \varphi(P)$ is stable at $(\succ, q; P)$. An allocation mechanism φ is **strategy-proof** if truth-telling is a weakly dominant strategy for all students, i.e., if for every student $i, \varphi(P)(i) R_i \varphi(P'_i, P_{-i})(i)$ for all P'_i and P_{-i}. Likewise, φ is **group strategy-proof** if there do not exist nonempty $J \subset I, P$ and $P'_J = (P'_i)_{i \in J}$ such that for all $i \in J, \varphi(P'_J, P_{-J})(i) R_i \varphi(P)(i)$, and for some $j \in J, \varphi(P'_J, P_{-J})(j) P_j \varphi(P)(j)$.

Given a problem $(\succ, q; P)$ and an assignment μ, suppose students in a nonempty set $J \subset I$ are removed together with their assigned seats. For each $s \in S$, let $q'_s \equiv q_s - |\{i \in J : \mu(i) = s\}|$. Let $I' \equiv I \backslash J$ and $S' \equiv \{s \in S : q'_s > 0\}$. Then the removal induces a unique **sub-priority-structure** $(\succ', q')$ and a **subproblem** $(\succ', q'; P'_{-J})$, where $\succ'$ (P'_{-J}, resp.) is the restriction of $\succ$ (P, resp.) to students in I' and schools in S', and q' is $(q'_s)_{s \in S'}$.

[6]For notational simplicity, we suppress I and S in the description of a problem.

2.2 Deferred acceptance algorithm

The **(student-proposing) deferred acceptance algorithm** (DA) is proposed in the classic work of Gale and Shapley (1962). Through DA, each priority structure $(\succ, q)$ induces an allocation mechanism which we denote by $DA^{\succ,q}$. For each preference profile P, $DA^{\succ,q}$ operates as follows:

Step 1. Each student applies to her favorite school. Each school tentatively accepts the best students according to its priority list up to its quota and rejects the rest.

Step $k, k \geq 2$. Each student rejected in the previous step applies to her next best school. Each school tentatively accepts the best students according to its priority list up to its quota and rejects the rest, among both new applicants and previously accepted students.

The algorithm stops when no student is rejected. DA assigns each student the last school that accepted her during the algorithm. Gale and Shapley (1962) show that DA produces the student-optimal stable assignment at $(\succ, q; P)$, which Pareto dominates any other stable assignment for the students. Also, due to Dubins and Freedman (1981) and Roth (1982), $DA^{\succ,q}$ is strategy-proof. However, $DA^{\succ,q}$ need not be Pareto efficient.

3 Top fair set

3.1 Definition

A set of students is called a top fair set if when all students apply to their favorite schools and all schools accept the best applicants up to their quotas, students in the set are always accepted, regardless of other students' preferences. Let $r_s(i) \equiv |\{j \in I : j \succ_s i\}| + 1$ denote the rank of student i at $\succ_s$.

Definition 1. *Fix a school choice problem $(\succ, q; P)$. A* ***top fair set (TFS)*** *is a nonempty set of students $T \subset I$ such that for every student $i \in T$ and her favorite school $s \in S \cup \{\emptyset\}$,*

$$r_s(i) - |\{i' \in T : i' \succ_s i, i' \text{ favors } s' \neq s\}| \leq q_s.$$

The inequality ensures that for each school s, students in T who favor s are all ranked among top-q_s by s, when students in T who favor other schools are excluded.[7] That is, the simultaneous assignments of students in T justify the fairness of each other. Equivalently, if T is a TFS, then after assigning to students in T, for each $i \in T$ and her favorite school s, unassigned students who have higher priorities than i at s do not outnumber the remaining seats at s.[8]

Example 1 below presents an illustration of TFS, and Example 2 below presents a simple school choice problem at which no TFS exists, although its DA assignment is Pareto efficient.

Example 1. Suppose $I = \{1, 2, 3, 4, 5\}, S = \{s_1, s_2, s_3\}, q_{s_1} = 2$, and $q_{s_2} = q_{s_3} = 3$. The priority lists are described in the table below. Let P be any preference profile such that student 1's favorite school is s_2, student 2's favorite is s_3, and students $3, 4, 5$'s favorite is

[7] To define TFS in the more general setting where each school s is associated with a choice function $C_s : 2^I \to 2^I$ as in Kojima and Manea (2010), this inequality should be replaced with $i \in C_s(\{i' \in T : i' \text{ favors } s\} \cup (I \backslash T))$, i.e., i will be accepted by s even if all students not in T also apply to s.

[8] That is, we can replace the inequality in the definition of TFS with $|\{j \notin T : j \succ_s i\}| \leq q_s - |\{i' \in T : i' \text{ favors } s\}|$.

s_1; students' favorite schools are described by the bottom row of the table.

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$
1	2	5
2	5	3
3	3	1
4	4	4
5	1	2
3,4,5	1	2

The problem $(\succ, q; P)$ has three TFS: $T_1 = \{1,2,3\}, T_2 = \{1,2,4\}$, and $T_3 = \{1,2,3,4\}$. To see this, take T_2 for example. It is easy to see that when students in T_2 apply to their respective favorite schools, no matter which schools students 3 and 5 apply to, students in T_2 will always be accepted.

Example 2. Suppose $I = \{1,2,3\}, S = \{s_1, s_2\}$, and $q_{s_1} = q_{s_2} = 1$. The priority lists are described in the left table below. Let P be any preference profile such that student 1's favorite school is s_2 and students 2 and 3's favorite is s_1; students' favorite schools are described by the bottom row of the left table.

$\succ_{s_1}$	$\succ_{s_2}$
1	3
2	1
3	2
2,3	1

P_1	P_2	P_3
s_2	s_1	s_1
$\emptyset$	s_2	s_2

The problem $(\succ, q; P)$ has no TFS. This is because since s_2 has only one seat, for student 1 to be in any TFS T, student 3–who can exclude student 1 from s_2–must also be in T. Likewise, for student 3 to be in T, so must student 1. However, assigning to both student 1 and student 3 will be excluded by student 2 at s_1. As a result, even for preference profile P at which $DA^{\succ,q}(P)$ is Pareto efficient (for example, P as described by the right table above), $(\succ, q; P)$ need not have TFS.

Next, we present some structural properties of TFS.

Proposition 1. *Fix a school choice problem $(\succ, q; P)$. The following properties hold:*

(i) *If both T and T' are TFS, then so is $T \cup T'$;*

(ii) *Suppose T is a TFS and $T' \subsetneq T$. If every student in T' is removed with a seat from her favorite school, then $T \backslash T'$ is still a TFS at the remaining subproblem.*

Due to part (i), if any TFS exists, then there must be a maximal TFS which is the union of all TFS. Part (ii) is straightforward from the definition of TFS; it reveals a form of consistency in how TFS makes assignments: the removal of a subset of students in T with their assignments won't affect the assignments of the rest of the students in T.

3.2 Finding TFS

For any school choice problem $(\succ, q; P)$, an **(iterative) elimination process** operates as follows:

Step 1. Let each student apply to her favorite school. Then let each school select the best applicants according to its priority list, up to its quota. Students who are not accepted by her favorite school are eliminated.

Step $t, t \geq 2$. For each $s \in S$, let every student who hasn't been eliminated still apply to her favorite school, and let all students who have ever been eliminated (by any school in any step before Step-t) apply to s. Then let s select the best applicants according to its priority list, up to its quota. Students who are not accepted by s are eliminated.

This process stops at the step when no new students are eliminated. For any given school choice problem, it always finds the maximal TFS, if any TFS exists.

Proposition 2. *The set of students who survive the elimination process, if nonempty, is the maximal TFS. If it is empty, then no TFS exists.*

To prove Proposition 2, we show by induction, that students eliminated in any step of the elimination process cannot be in any TFS. Let's illustrate the elimination process with previous examples.

Example 3. In Example 1, only student 5 is rejected (by s_1) in Step-1 of the elimination process. Therefore, in Step-2, student 5 applies to each school to see whether any accepted student can be newly eliminated; the answer is no and the elimination process stops. The set of students who survive is $\{1, 2, 3, 4\}$.

In Example 2, only student 3 is rejected (by s_1) in Step-1 of the elimination process. In Step-2, student 3 applies to each school, and student 1 is eliminated from s_2. In Step-3, all the eliminated students, i.e., student 3 and 1, apply to each school, and student 2 is now eliminated from s_1. By now, all students have been eliminated and the elimination process stops.

4 Main result

Fix any given priority structure $(\succ, q)$, we define the $TFS^{\succ,q}$ algorithm. For each preference profile P of students, $TFS^{\succ,q}$ operates as follows:

Step 1. If the problem $(\succ, q; P)$ has no TFS, stop. Otherwise, find a TFS, assign each student in it a seat at her favorite school, and then remove these students with their assigned seats.

Step $t, t \geq 2$. Consider the subproblem induced by the removal in previous steps. If it has no TFS, stop. Otherwise, find a TFS, assign each student in it a seat at her favorite school, and then remove these students with their assigned seats.

We did not specify how to select a TFS when multiple exist in some step of the algorithm. As our next result shows, due to Proposition 1, such selection does not affect the outcome of the algorithm. A default option is to choose the maximal TFS in each step of $TFS^{\succ,q}(P)$, which can be found through the elimination process (Proposition 2).

Proposition 3. *For any school choice problem, the outcome of the TFS algorithm is independent of the choice of TFS in each step of the algorithm.*

Also, the outcome of the TFS algorithm, $TFS^{\succ,q}(P)$, may only be a partial assignment, because the algorithm stops whenever TFS does not exist.

Definition 2. *A priority structure $(\succ, q)$ is* ***TFS-solvable*** *if at any P, $TFS^{\succ,q}$ produces a complete assignment.*

Our main result shows the equivalence between TFS-solvability and the efficiency of the DA mechanism.

Theorem 1. *A priority structure $(\succ, q)$ is TFS-solvable if and only if $DA^{\succ,q}$ is a Pareto efficient mechanism. Moreover, if $(\succ, q)$ is TFS-solvable, then $TFS^{\succ,q}(P) = DA^{\succ,q}(P), \forall P$.*

Therefore, the assignment of any Pareto efficient DA mechanism can always be decomposed into a sequence of TFS assignments. In this sense, Pareto efficient DA mechanisms share a similar iterative removal structure with other known group strategy-proof and Pareto efficient mechanisms.[9]

The set of priority structures that induce Pareto efficient DA mechanisms have been characterized by Ergin (2002)'s acyclicity condition.[10] A priority structure $(\succ, q)$ is **acyclic** if there do not exist distinct schools s_1, s_2 and distinct students i, j, k such that: (i) $i \succ_{s_1} j \succ_{s_1} k \succ_{s_2} i$; and (ii) there exist disjoint sets of students $I_{s_1}, I_{s_2} \subset I \backslash \{i, j, k\}$ such that $I_{s_1} \succ_{s_1} j, I_{s_2} \succ_{s_2} i, |I_{s_1}| = q_{s_1} - 1$, and $|I_{s_2}| = q_{s_2} - 1$. For any $(\succ, q), DA^{\succ,q}$ is **consistent** if for any P, after removing students in any set $J \subset I$ with their assignments at $DA^{\succ,q}(P)$, at the subproblem $(\succ', q'; P'_{-J})$ induced by the removal, $DA^{\succ',q'}(P'_{-J})$ assigns students in $I \backslash J$ the same assignments as in $DA^{\succ,q}(P)$.

Theorem 2 (Ergin, 2002)**.** *For any $(\succ, q)$, the following are equivalent:*

(i) *$DA^{\succ,q}$ is Pareto efficient;*

(ii) *$DA^{\succ,q}$ is group strategy-proof;*

(iii) *$DA^{\succ,q}$ is consistent;*

(iv) *$(\succ, q)$ is acyclic.*

[9] When $q_s = 1, \forall s$, these mechanisms have been characterized by Pycia and Ünver (2017) as trading cycles mechanisms, which generalize Pápai (2000)'s hierarchical exchange rules by introducing a new form of ownership called brokerage. Pycia and Ünver (2011) extend trading cycles mechanisms to school choice with general capacities.

[10] Ergin (2002) also provides a characterization of acyclicity by showing that the priority lists of any pair of schools should be similar in positions lower than the sum of their quotas.

The following observation follows immediately from Theorem 1 and Theorem 2.

Corollary 1. *A priority structure $(\succ, q)$ is TFS-solvable if and only if it is acyclic.*

Suppose $(\succ, q)$ is acyclic. From Ergin's equivalence results, $DA^{\succ,q}$ (and equivalently, $TFS^{\succ,q}$) is Pareto efficient, group strategy-proof, and consistent.[11] We argue that these properties of $DA^{\succ,q}$ become more intuitive, when we look at its TFS decomposition. First, Pareto efficiency and group strategy-proofness are due to assigning students their favorite schools and the iterative structure in the TFS algorithm, the same reasons that the TTC-based mechanisms satisfy these properties. Second, consistency is due to part (ii) of Proposition 1: the removal of a subset of students in T with their assignments won't affect the assignments of the rest of the students in T.

The main difficulty in proving Theorem 1 lies in proving the following lemma.

Lemma 1. *If $DA^{\succ,q}$ is Pareto efficient, then for any P, a TFS exists.*

To prove this lemma, it is sufficient to show that if $DA^{\succ,q}$ is Pareto efficient, then at any P, a nonempty set of students survive the elimination process. The key intuition is that if any student k is eliminated from school s in the first step and k is able to eliminate some i accepted by s' in the second step, then i or any student with lower priority than i at s' cannot help eliminate any first-step students accepted by s. This is because otherwise there will be a rejection cycle in the DA procedure of certain preference profile: s rejects k and k's next application to s' leads to the rejection of some i, whose application to s in turn excludes some student accepted by s. The existence of such rejection cycle is in conflict with the efficiency of DA. More generally, if for all $2 \leq l \leq L$, students eliminated by s_l can help eliminate students at s_{l-1}, then students eliminated by s_1 cannot help eliminate

[11] Kojima and Manea (2010) show that for DA mechanisms under acceptant substitutable priority, Pareto efficiency, Maskin monotonicity, and group strategy-proofness are equivalent. The equivalence between these axioms and consistency is extended to acceptant substitutable priority by Klijn (2011).

students at s_L. Therefore, there must be a school whose first-step accepted students all survive the second-step elimination. By induction, this holds true for all steps.

5 Conclusion

We propose a new concept, top fair set, to identify which students should be assigned their favorite school. When the priority structure is acyclic, iterative assignment of top fair sets ensures both Pareto efficiency and group strategy-proofness. Our main result shows that each Pareto efficient DA mechanism can be decomposed into a sequence of TFS. If a priority structure is not acyclic, it is not TFS-solvable. We do not yet know how to properly extend the TFS algorithm to such priority structures. A natural way is to assign and remove TFS when it exists, and when no TFS exists, assign and remove TTC instead. However, such an algorithm is not strategy-proof. For the sake of simplicity, we restrict attention to priority structures that consist of priority lists. When schools have acceptant substitutable priority, the DA mechanism is well-defined and is still student-optimally stable. TFS is also well-defined for such priority structures, and it is not difficult to extend the equivalence between TFS-solvability and the Pareto efficiency of the DA mechanism.

A Appendix

A.1 Proof of Proposition 1

Proof. For part (i), suppose both T and T' are TFS. For each s that is favored by some student in $T \cup T'$, let $i \in T \cup T'$ be the student with the lowest priority at s among students in $T \cup T'$ who favor s. Without loss of generality, assume $i \in T$. Since T is a TFS, $|\{j \notin T :$

$j \succ_s i\}| \leq q_s - |\{i' \in T : i'$ favors $s\}|$. For each $j' \in T' \backslash T$ who favors s, by assumption, $j' \succ_s i$. Therefore, from T to $T \cup T'$, each addition of a student $j' \in T' \backslash T$ who favors s reduces a student not in T who has higher priority than i at s. As a result, $|\{j \notin T \cup T' : j \succ_s i\}| + |\{i' \in T \cup T' : i'$ favors $s\}| = |\{j \notin T : j \succ_s i\}| + |\{i' \in T : i'$ favors $s\}| \leq q_s$. Moreover, this inequality is also satisfied by any other student in $T \cup T'$ who favors s, because by assumption such a student is ranked higher than i at s.

For part (ii), suppose students in T' are removed together with their TFS assignments. Then for each school s such that a number of seats of it are removed, in the subproblem, $|\{i' \in T : i'$ favors $s\}|$ is reduced by the same number. Therefore, in this subproblem, for each $i \in T \backslash T'$ who favors this school s, $|\{j \notin T : j \succ_s i\}| \leq q'_s - |\{i' \in T \backslash T' : i'$ favors $s\}|$ still holds. □

A.2 Proof of Proposition 2

Proof. We first introduce some notations. For each school s, denote by $R_t(s)$ the set of students who favor s and are accepted by s in the first $t-1$ steps, but are (newly) eliminated by it in Step-t of the elimination process. Likewise, denote by $A_t(s)$ the set of students who favor s and are accepted by s in the first t steps. Then $R_t \equiv \cup_{s \in S} R_t(s)$ is the set of all newly eliminated students in Step-t, and $A_t \equiv \cup_{s \in S} A_t(s)$ is the set of students who survive the first t steps of elimination. By definition, $A_t = I \backslash (R_1 \cup \cdots \cup R_t)$. If the elimination process stops at Step-$\bar{t}$, for all $t \geq \bar{t}$ and all s, let $A_t(s) = A_{\bar{t}}(s)$ and $R_t(s) = \emptyset$.

Suppose a nonempty set of students who survive the elimination process; denote it by M . Then M is a TFS, as no matter what preferences students in $I \backslash M$ have, students in M are always accepted by their favorite schools. We next show that M is the maximal TFS. Fix an arbitrary TFS T, we only need to show that, by induction, it contains no student who has ever been eliminated along the process.

Let us start with $R_1 \cap T = \emptyset$. Suppose not, then there is a student i eliminated from her favorite school s in Step-1 and $i \in T$. Note that $A_1(s) \succ_s i$ and $|A_1(s)| = q_s$. Since every student in $A_1(s)$ either belongs to $\{j \notin T : j \succ_s i\}$ or $\{i' \in T : i' \text{ favors } s\}$, $|\{j \notin T : j \succ_s i\}| + |\{i' \in T : i' \text{ favors } s\}| \geq q_s + 1$. A contradiction to $i \in T$.

Suppose for some natural number $k < \bar{t}$, $R_t \cap T = \emptyset$ for all $1 \leq t \leq k$. We now prove $R_{k+1} \cap T = \emptyset$. If not, then there exists a student $i \in R_{k+1}(s) \cap T$, where s is her favorite school. Then i is not selected by her favorite school s among students in $A_k(s) \cup R_1 \cup \cdots \cup R_k$. There are at least $q_s - |A_{k+1}(s)|$ students in $R_1 \cup \cdots \cup R_k$ who have higher priority than i at s. Due to the induction hypothesis, these students in $R_1 \cup \cdots \cup R_k$ belong to $\{j \notin T : j \succ_s i\}$. It is also clear that $A_{k+1}(s) \succ_s i$. As a result, every student in $A_{k+1}(s)$ either belongs to $\{j \notin T : j \succ_s i\}$ or $\{i' \in T : i' \text{ favors } s\}$. Therefore, $|\{j \notin T : j \succ_s i\}| + |\{i' \in T : i' \text{ favors } s\}| \geq q_s + 1$. A contradiction to $i \in T$. □

A.3 Proof of Proposition 3

Proof. Fix a school choice problem $(\succ, q; P)$. Let $\{M_1, M_2, \ldots, M_n\}$ be the sequence of *maximal* top fair sets in the TFS algorithm, and denote by μ the outcome (a partial assignment or assignment). Fix another arbitrary TFS process, denoted the sequence of top fair sets by $\{T_1, T_2, \ldots, T_m\}$ and the final outcome by ν. Our goal is to show that $\cup_{k=1}^{m} T_k = \cup_{l=1}^{n} M_l$, and moreover $\mu = \nu$ when restricted to this union of subsets. For every subset of students $J \subset I$, we write $\mu|_J = \nu|_J$ if for every $j \in J, \mu(j) = \nu(j)$.

We proceed to prove this result into the following three steps.

Step I. *$M_1 \subset \cup_{k=1}^{m} T_k$, and $\nu|_{M_1} = \mu|_{M_1}$.*

The statement holds if $M_1 = T_1$. In fact, it is clear that $T_1 \subset M_1$ because M_1 is the maximal TFS in the original problem. Note that $\mu|_{T_1} = \nu|_{T_1}$ as every student in T_1 gets her favorite school. Suppose that $T_1 \subsetneq M_1$. In the remaining subproblem after removing T_1, both T_2 and $M_1 \backslash T_1$ are TFS. Therefore, for every student $j \in (M_1 \backslash T_1) \cap T_2 = M_1 \cap T_2$, her favorite school in the original problem is still guaranteed, as a result, $\mu|_{M_1 \cap T_2} = \nu|_{M_1 \cap T_2}$. Similarly, we can show by induction that, for every $k = 2, \ldots, m$, if $M_1 \backslash (T_1 \cup T_2 \cup \cdots \cup T_k) \neq \emptyset$, it is a TFS in the remaining subproblem when $T_1, T_2, \ldots, T_k$ are removed sequentially, and then $\mu|_{M_1 \cap T_{k+1}} = \nu|_{M_1 \cap T_{k+1}}$. Because the TFS procedure $\{T_1, T_2, \ldots, T_m\}$ terminates when there is no TFS, it must be the case that $M_1 \backslash (T_1 \cup T_2 \cup \cdots \cup T_m) = \emptyset$. Therefore, $M_1 \subset \cup_{k=1}^m T_k$ and then $\nu|_{M_1} = \mu|_{M_1}$.

Step II. *The sequence of subsets, $\{M_1, T_2 \backslash M_1, T_3 \backslash M_1, \ldots, T_m \backslash M_1\}$, consists of top fair sets for a TFS procedure in the original problem (if for any k, $T_k \backslash M_1 = \emptyset$, just drop it from the sequence). Moreover, this TFS procedure yields the outcome ν.*

The statement is true if $T_1 = M_1$. Suppose that $T_1 \subsetneq M_1$. From Step I, every student in the maximal TFS M_1 is assigned her favorite school and leaves the market. If $T_2 \backslash M_1$ is nonempty, it is a TFS in the remaining subproblem after removing M_1. In fact, the removal of M_1 can be decomposed into two sub-steps, removing T_1 first, then removing $M_1 \backslash T_1$, as a result, $T_2 \backslash (M_1 \backslash T_1) = T_2 \backslash M_1$ remains a TFS in the remaining subproblem after removing M_1. Meanwhile, students in $T_2 \backslash M_1$ still get their favorite schools in the remaining subproblem after the removal of T_1. Similarly, one can show by induction that $\{M_1, T_2 \backslash M_1, T_3 \backslash M_1, \ldots, T_m \backslash M_1\}$ (if for any k, $T_k \backslash M_1 = \emptyset$, just drop it from the sequence) consists of a sequence of top fair sets for a TFS procedure. Moreover, every student in $T_{k+1} \backslash M_1$ is assigned her favorite school as in ν in the remaining subproblem after the removal of $T_1, T_2, \ldots, T_k$. Therefore, the TFS procedure $\{M_1, T_2 \backslash M_1, \ldots, T_k \backslash M_1\}$ yields the same outcome ν.

Step III. We have proved in Step I that $M_1 \subset \cup_{k=1}^{m} T_k$, and $\nu|_{M_1} = \mu|_{M_1}$. Remove M_1 and their assignments as in μ from the market. In the remaining subproblem, it is clear that $\{M_2, \ldots, M_l\}$ is the TFS procedure in which the maximal TFS is removed at each step, and each student is assigned the same school as in μ. From Step II, $\{T_2 \backslash M_1, T_3 \backslash M_1, \ldots, T_m \backslash M_1\}$ (if some subset is empty, drop it) is another TFS procedure in this subproblem, and each student is assigned the same school as in ν. It is clear that, $T_2 \backslash M_1 \subset M_2$, thus $T_2 \subset M_1 \cup M_2$, because M_2 is the maximal TFS in this subproblem. Next, apply the argument in Step I for these two TFS procedures in this subproblem, $M_2 \subset \cup_{k=2}^{m}(T_k \backslash M_1)$, thus $M_2 \subset \cup_{k=1}^{m} T_k$, and $\nu|_{M_2} = \mu|_{M_2}$. Finally, this argument can be applied by induction to verify that $M_l \subset \cup_{k=l}^{m} T_k$, and $\nu|_{M_l} = \mu|_{M_l}$ for all $l = 1, \ldots, n$. We therefore proved $\cup_{k=1}^{m} T_k = \cup_{l=1}^{n} M_l$, and μ and ν coincide when restricted on this set. $\square$

A.4 Proof of Lemma 1

Proof. Suppose $DA^{\succ,q}$ is Pareto efficient and P is a given preference profile. Consider the elimination process associated with the problem $(\succ, q; P)$. Recall that from the proof of Proposition 2, $R_t(s)$ denotes the set of students who favor s but are (newly) eliminated by it in Step-t of the elimination process, and $A_t(s)$ denotes the set of students who favor s and are (still) accepted by s in Step-t. Also, $R_t \equiv \cup_{s \in S} R_t(s)$ and $A_t \equiv \cup_{s \in S} A_t(s)$.

Claim 1. *Suppose $DA^{\succ,q}$ is Pareto efficient and $s_1, s_2, \ldots, s_L$ is a collection of schools. If for every $2 \leq l \leq L$, $A_1(s_{l-1}) \succ_{s_{l-1}} R_1(s_l)$ does not hold, then $A_1(s_L) \succ_{s_L} R_1(s_1)$.*

When $L = 2$, this claim simply says that if some student eliminated at s_2 in the first step is able to eliminate first-step accepted students at s_1 in the second step, i.e., if not $A_1(s_1) \succ_{s_1} R_1(s_2)$, then students in $R_1(s_1)$ cannot eliminate any student in $A_1(s_2)$, i.e., $A_1(s_2) \succ_{s_2} R_1(s_1)$.

Proof. Assume for every $2 \le l \le L$, $A_1(s_{l-1}) \succ_{s_{l-1}} R_1(s_l)$ does not hold. Suppose instead $A_1(s_L) \succ_{s_L} R_1(s_1)$ also does not hold. Then for each $2 \le l \le L$, there exists $k_l \in R_1(s_l)$ and $i_{l-1} \in A_1(s_{l-1})$ such that $k_l \succ_{s_{l-1}} i_{l-1}$. In addition, there exists $k_1 \in R_1(s_1)$ and $i_L \in A_1(s_L)$ such that $k_1 \succ_{s_L} i_L$. Construct a preference profile P' such that for all $k_l, 2 \le l \le L$, her new preference is $P'_{k_l} : s_l, s_{l-1}, \emptyset$, and $P'_{k_1} : s_1, s_L, \emptyset$; for all $j \in A_1(s)$ for some s (except students i_l, $1 \le l \le L$), her new preference is $P'_j : s, \emptyset$; for all the i_l's and all other students, assume they find all schools unacceptable.

We will see that $DA(\succ, q; P')$ is not Pareto efficient, which contradicts with the assumption that $DA^{\succ,q}$ is Pareto efficient. In the DA procedure of $(\succ, q; P')$, s_L accepts students in $A_1(s_L)$ and rejects k_L. After that, k_L applies to s_{L-1} and s_{L-1} rejects k_{L-1}. Eventually, k_1 applies to s_L and since $k_1 \succ_{s_L} i_L$ for some $i_L \in A_1(s_L)$, s_L will reject some student in $A_1(s_L)$. After that, DA stops. The DA assignment is not Pareto efficient because letting students $k_1, \ldots, k_L$ exchange their assignments leads to a Pareto improvement. □

Claim 2. *Suppose $DA^{\succ,q}$ is Pareto efficient. For all preference profile P, $A_2 \ne \emptyset$.*

Proof. It is sufficient to show that for some school s with $A_1(s) \ne \emptyset$, $R_2(s) = \emptyset$. Suppose not. Then for every s_0 such that $A_1(s_0) \ne \emptyset$, there exists s_1 such that $A_1(s_0) \succ_{s_0} R_1(s_1)$ fails to hold. Accordingly, there exists s_2 such that $A_1(s_1) \succ_{s_1} R_1(s_2)$ fails to hold, and due to Claim 1, $s_2 \ne s_0$. By induction and iteratively applying Claim 1, for any $L \ge 1$, there exist heterogeneous schools $s_1, s_2, \ldots, s_L$ such that for each $1 \le l \le L$, $A_1(s_{l-1}) \succ_{s_{l-1}} R_1(s_l)$ fails to hold. However, since there are only finitely many schools, this is not possible. □

Claim 3. *Suppose $DA^{\succ,q}$ is Pareto efficient and $t \ge 2$. If for all preference profile P, $A_t \ne \emptyset$, then for all P, $A_{t+1} \ne \emptyset$.*

Proof. Suppose for a given t, for all preference profile P, during the elimination process of $(\succ, q; P)$, $A_t \ne \emptyset$. For any given P, we can reduce its $(t+1)$-th step elimination to be the

t-th step elimination of a modified preference profile.

Consider students in R_2 at the elimination process under $(\succ, q; P)$. If $i \in R_2(s_0)$ for any s_0, then due to our arguments in proving Claim 2, there exist heterogeneous schools $s_1, s_2, \ldots, s_L$ such that there exists $k \in R_1(s_1)$ satisfying $k \succ_{s_0} i$, and for each $1 \leq l \leq L$, $A_1(s_{l-1}) \succ_{s_{l-1}} R_1(s_l)$ fails to hold, while for all s, $A_1(s_L) \succ_{s_L} R_1(s)$. That is, no student in $A_1(s_L)$ will be eliminated in the second step of elimination; $R_2(s_L) = \emptyset$.

In addition, if $i \in R_2(s_0)$ and there exists $k \in R_1(s_1)$ such that $k \succ_{s_0} i$, then $A_1(s_1) \succ_{s_1} i$. This is because otherwise due to the same argument as in the proof of Claim 1, DA will be inefficient: under certain preferences, during DA, k is rejected from s_1, and then she applies to s_0, causing i to be rejected from s_0, and i then applies to s_1, causing some student in $A_1(s_1)$ to be rejected.

By induction, $A_1(s_L) \succ_{s_L} i$. Then we know that if any $i \in R_2(s_0)$ lists s_L as her favorite school, then i will be rejected in the first step of the elimination process. For every $i \in R_2$, we can find such a school s_L and modify P_i such that i lists s_L as her favorite school. Denote the new preference profile as P' and the step-$\tilde{t}$ eliminated students under $(\succ, q; P')$ as R'_t. Then $R'_1 = R_2 \cup R_1$, because all students in R_2 will now be eliminated in step-1 under P'.

As a result, for any $\tilde{t} \geq 2, R'_{\tilde{t}}(s) = R_{\tilde{t}+1}(s)$ and $A'_{\tilde{t}}(s) = A_{\tilde{t}+1}(s), \forall s$. By assumption, for all preference profiles, $A_t \neq \emptyset$. Apply this to P'. Then $A'_t = A_{t+1} \neq \emptyset$. $\square$

The proof of Lemma 1 then follows directly from Claim 2 and 3 by induction on t. $\square$

A.5 Proof of Theorem 1

Proof. It is easy to see that if $(\succ, q)$ is TFS-solvable, then for any $P, TFS^{\succ,q}(P)$ is always stable and Pareto efficient at $(\succ, q; P)$. We also know that $DA^{\succ,q}(P)$ is always optimally stable at $(\succ, q; P)$. Therefore, if $(\succ, q)$ is TFS-solvable, then $TFS^{\succ,q}(P) = DA^{\succ,q}(P), \forall P$, and $DA^{\succ,q}$ is Pareto efficient.

Therefore, we only need to show that if $DA^{\succ,q}$ is Pareto efficient, then $(\succ, q)$ is TFS-solvable. For necessity, suppose $DA^{\succ,q}$ is efficient. Fix any preference profile P. Due to Lemma 1, a TFS exists. Then $TFS^{\succ,q}$ can find a TFS T, assign students in it with seats at their favorite schools, and then remove them.

Let $(\succ', q')$ be the sub-priority-structure induced by the assignment and removal of T. Then $DA^{\succ',q'}$ is also Pareto efficient mechanism. This is because students in T must have been assigned and removed with their assignment at $DA(\succ, q; P)$. Due to the consistency of efficient DA mechanisms shown by Theorem 1 of Ergin (2002), at the remaining subproblem, DA produces the same Pareto efficient assignment for students in $I \backslash T$ as $DA(\succ, q; P)$. Furthermore, this is true regardless of the preferences of students in $I \backslash T$. Therefore, $DA^{\succ',q'}$ is efficient.

By induction, after the removal of any TFS, we can find another TFS at the remaining subproblem. Such iterative removal of TFS stops only when all students are removed. Therefore, $DA^{\succ,q}$ is Pareto efficient implies that $(\succ, q)$ is TFS-solvable. □

References

A. Abdulkadiroğlu, Y-K. Che, P. Pathak, A. Roth, O. Tercieux (2019), Efficiency, justified envy, and incentives in priority-based matching, *American Economic Review: Insights,* forthcoming.

A. Abdulkadiroğlu, T. Sönmez (2003), School choice: A mechanism design approach, *American Economic Review*, 93, 729-747.

M. Balinski, T. Sönmez (1999), A tale of two mechanisms: Student placement, *Journal of Economic Theory*, 84, 73-94.

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

L. Ehlers, B. Klaus (2006), Efficient priority rules, *Games and Economic Behavior*, 55, 372-384.

H. Ergin (2002), Efficient resource allocation on the basis of priorities, *Econometrica*, 70, 2489-2497.

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

G. Haeringer, F. Klijn (2009), Constrained school choice, *Journal of Economic Theory*, 144, 1921-1947.

R. Hakimov, O. Kesten (2018), The equitable top trading cycles mechanism for school choice, *International Economic Review,* 59, 2219-2258.

O. Kesten (2004), Student placement to public school in US: Two new solutions, *working paper*.

O. Kesten (2006), On two competing mechanisms for priority-based allocation problems, *Journal of Economic Theory*, 127, 155-171.

F. Klijn (2011), On the consistency of deferred acceptance when priorities are acceptant substitutable, *Mathematical Social Sciences*, 62, 101-103.

F. Kojima (2011), Robust stability in matching markets, *Theoretical Economics*, 6, 257-267.

F. Kojima, M. Manea (2010), Axioms for deferred acceptance, *Econometrica*, 78, 633-653.

T. Kumano (2009), Efficient resource allocation under acceptant substitutable priorities, *working paper*.

J. Ma (1994), Strategy-proofness and the strict core in a market with strict indivisibilities, *International Journal of Game Theory*, 23, 75-83.

T. Morrill (2015), Two simple variations of top trading cycles, *Economic Theory*, 1, 123-140.

S. Pápai (2000), Strategyproof assignment by hierarchical exchange, *Econometrica*, 68, 1403-1433.

M. Pycia, U. Ünver (2011), Trading cycles for school choice, *working paper*.

M. Pycia, U. Ünver (2017), Incentive compatible allocation and exchange of discrete resources, *Theoretical Economics*, 12, 287-329.

A. Roth (1982), The economics of matching: stability and incentives, *Mathematics of Operations Research*, 7, 617-628.

A. Roth, M. Sotomayor (1990), Two-sided matching: A study in game-theoretic modeling and analysis, Cambridge University Press.

M. A. Satterthwaite, H. Sonnenschein (1981), Strategy-proof allocation mechanisms at differentiable points, *Review of Economic Studies*, 48, 587-597.

L. Shapley, H. Scarf (1974), On cores and indivisibility, *Journal of Mathematical Economics*, 1, 23-37.

L. G. Svensson (1994), Queue allocation of indivisible goods, *Social Choice and Welfare*, 11, 323-330.