Improving the Deferred Acceptance with Minimal Compromise*

Mustafa Oğuz Afacan† Umut Dur‡ A. Arda Gitmez§
Özgür Yılmaz¶

June 27, 2022

Abstract

In school choice problems, the motivation for students' welfare (*efficiency*) is restrained by concerns to respect schools' priorities (*fairness*). Even the best matching in terms of welfare among all *fair* matchings ($SOSM$) is in general *inefficient*. Moreover, any mechanism that improves welfare over the $SOSM$ is manipulable by the students. First, we characterize the "least manipulable" mechanisms in this class: *upper-manipulation-proofness* ensures that no student is better off through strategic manipulation over the objects that are better than their assigned school. Second, we use the notion that a matching is *less unfair* if it yields a smaller set of students whose priorities are violated, and define *minimal unfairness* accordingly. We then show that the *Efficiency Adjusted Deferred Acceptance* ($EADA$) mechanism is *minimally unfair* in the class of *efficient* and *upper-manipulation-proof* mechanisms.When the objective is to improve students' welfare over the $SOSM$, this characterization implies an important insight on the frontier of the main axioms in school choice.

*We thank Aram Grigoryan for his useful comments.
†Faculty of Art and Social Sciences, Sabancı University, Istanbul, Turkey. E-mail: mafacan@sabanciuniv.edu
‡Department of Economics, NC State University, Raleigh, NC, USA. E-mail: umutdur@gmail.com
§Department of Economics, Bilkent University, Ankara, Turkey. E-mail: arda.gitmez@bilkent.edu.tr
¶Department of Economics, Koç University, Istanbul, Turkey. E-mail: ozyilmaz@ku.edu.tr

1 Introduction

Allocating school seats to students is an important problem for school districts. In a *school choice problem*, students submit their preferences over schools and the central placement authority allocates seats to students based on submitted preferences and schools' priorities over students. The role of priorities is crucial: respecting priorities (*fairness*) has been the primary desideratum for the school choice problem. On the other hand, the main motivation for adopting a *market design* approach to school choice has been to improve students' welfare (*efficiency*). After all, the primary appeal of this approach is providing students with the opportunity to choose their school. Unfortunately, *efficiency* is incompatible with *fairness* (Roth, 1982; Balinski and Sönmez, 1999). This paper contributes to the understanding of the compromise between these two properties.

The tension between *fairness* and *efficiency* has compelled the policy makers to choose between the two conflicting objectives. In several large matching markets, including the New York City and Boston, *fairness* concerns prevailed, and the school districts have been implementing the well-known *Deferred Acceptance* (DA) algorithm (Gale and Shapley, 1962). The DA satisfies *justified-envy-freeness*, the canonical criterion for *fairness*: for each school s, there should not be a student who prefers s to their assigned school and another student assigned to s with a lower priority at s. The DA is favored for two other reasons. First, it finds the *student-optimal stable matching* (SOSM); that is, the best matching in terms of students' welfare among all *justified-envy-free* matchings (Gale and Shapley, 1962; Balinski and Sönmez, 1999). Second, it is immune to strategic manipulations by students, a property known as *strategy-proofness* (Dubins and Freedman, 1981; Roth, 1982).

A more recent approach to this tension is to prioritize *efficiency* and to compare mechanisms in terms of *fairness*. While the DA produces the "most efficient" matching among the set of *fair* matchings, this recent approach focuses on the "minimally unfair" matching among the set of *efficient* matchings. The perspective of the current work is the latter. We study the class of *efficient* mechanisms that improve students' welfare over the DA.

Welfare improvement over the DA requires compromising *strategy-proofness* (Abdulkadiroğlu et al., 2009; Kesten, 2010). Thus, the cost of welfare gains is not only *fairness* but also immunity to strategic manipulation. This negative result raises a challenging question: within the class of mechanisms that are *efficient* and improve welfare over the DA, what is the mechanism that is *least vulnerable to strategic manipulations* and *minimally unfair*? Our answer to this question is the *Efficiency-Adjusted Deferred Acceptance* ($EADA$) mechanism (Kesten, 2010). It is known that the $EADA$ is *efficient* and it improves welfare over the DA. We show that the $EADA$ is also distinguished in terms of *fairness* and immunity to strategic manipulation. To formulate our characterization result, we need to clarify the two associated concepts: (1) *least vulnerable to strategic manipulations*, and (2) *minimally unfair*.

The first conceptualization requires the decomposition of *strategy-proofness* into two weaker notions: *upper-manipulation-proofness* and *lower-manipulation-proofness*. *Upper - manipulation - proofness* (*Lower - manipulation - proofness*) ensures that no student is better off through strategic manipulation over the objects that are better (worse) than their assigned school. We show that a mechanism is *strategy-proof* if and only if it is *upper-manipulation-proof* and *lower-manipulation-proof* (Theorem 1). Moreover, no mechanism that improves welfare over the DA can be *lower-manipulation-proof* (Proposition 1). Thus, among the mechanisms that improve welfare over the DA, *upper-manipulation-proof* mechanisms, if any, are strategically the most robust. We show that the $EADA$ is *upper-manipulation-proof* (Proposition 3).

The second conceptualization is based on the following *fairness* comparison: a matching is *less unfair* if it yields a smaller set of students whose priorities are violated. Also, a matching is *equally unfair* if it yields the same set of students whose priorities are violated. A mechanism is *minimally unfair* in a class of mechanisms if there is no other mechanism in that class that produces *equally* or *less unfair* matching for every problem and a *less unfair* matching for at least one problem.

We restrict attention to *efficient* and *upper-manipulation-proof* mechanisms, and characterize the $EADA$ as a *minimally unfair* mechanism within this class (Theorem 2). A direct implication of this result is that the *EADA* is *minimally unfair* within the class of *upper-manipulation-proof* and *efficient* mechanisms improving over the *DA*. This implies an important insight on the frontier of basic principles in school choice: in securing all possible welfare improvements over the *DA* outcome, the *EADA* guarantees the least possible departure both from *fairness* and also from *strategy-proofness*. Thus, our results constitute a strong argument for the *EADA* as a solution to the conflict between *efficiency* and *fairness*. In Section 5.3, we also discuss the merits of our notions of vulnerability to strategic manipulations and fairness comparison beyond the theoretical interest.

Perhaps surprisingly, *upper-manipulation-proofness* is necessary for this characterization of the *EADA*. That is, the *EADA* is **not** *minimally unfair* in the class of *efficient* mechanisms that improve welfare over the *DA* (Doğan and Ehlers, 2021a). On the other hand, the *EADA* is *minimally unfair* with respect to an alternative notion of fairness comparison in the class of *efficient* mechanisms (Kwon and Shorrer, 2020).[1] We show that our *fairness* comparison is independent from this alternative notion (Proposition 2). Thus, the main results in these two works are independent from each other.

Related Literature

An implication of the incompatibility between *fairness* and *efficiency* is that the *DA* is not *Pareto efficient* (Balinski and Sönmez, 1999; Abdulkadiroğlu and Sönmez, 2003). Moreover, the inefficiency of the *DA* can be potentially severe (Kesten, 2010). Empirical findings on the NYC high school match give support to this theoretical possibility (Abdulkadiroğlu et al., 2009). There are three preeminent approaches to restoring *efficiency*.

The first approach is to weaken *justified-envy-freeness*. Such a weaker notion is *reasonable justified-envy-freeness*: a matching is *reasonably justified-envy-free* if whenever the priority of student i is violated at school s, there does not exist a *justified-envy-free* matching in which i is assigned to s (the working paper version of Kesten (2010)). Another weaker notion is $\alpha-$*equitability*: a matching with a priority violation is not considered as 'unfair'

[1] We also show that this characterization of the *EADA* can be strengthened by adopting a more demanding notion of being *more fair* (Proposition 5 in Appendix D).

if a student's objection to that priority violation is counter-objected by another student (Alcalde and Romero-Medina, 2017). Also, a matching is *essentially stable* if any priority-based claim initiates a chain of reassignments that results in the initial claimant losing the seat, and the $EADA$ selects such a matching (Troyan et al., 2020). It is also possible to formulate *fairness* of a set of matchings: a set of matchings is *legal* if and only if any matching outside the set has *justified-envy* with some matching in the set and no two matchings inside the set block each other via *justified-envy*. Under commonly assumed conditions on priorities, there always exists a unique *legal* set of matchings, and this set shares similarities with the set of *justified-envy-free* matchings: (i) it is a lattice and (ii) it satisfies the rural hospitals theorem. The *student-optimal legal* matching in this set is *efficient* and it is the outcome of a generalized version of the $EADA$ (Ehlers and Morrill, 2020).

The second approach is to eliminate (at least some of) the constraints due to *justified-envy-freeness*. The school district can ask students whether they consent their priorities to be violated and then allocates the seats at schools by using this additional data on students' consent decisions (Kesten, 2010). When priority violations are consented, the constraints due to *justified-envy-freeness* are relaxed and students' welfare can be improved. A necessary condition for this mechanism to work is that students should have incentives to consent for the violation of their priorities, that is, they should not be hurt by consenting. A mechanism motivated by this idea is the *Efficiency Adjusted Deferred Acceptance* ($EADA$) mechanism, which gives students incentives to consent (Kesten, 2010).

The idea of relaxing *fairness* through a set of *allowable priority violations* is generalized and formulated as *partial fairness* (Dur et al., 2019). A class of mechanisms for this extended problem with priority violations is the *Student Exchange under Partial Fairness* $SEPF$, where each mechanism in this class gives a *partially fair* matching that is not *Pareto dominated* by another *partially fair* matching, that is a *constrained efficient* matching in the class of *partially fair* matchings (Dur et al., 2019). A particular mechanism in the $SEPF$ class stands out: the *Top Priority* (TP) mechanism, which is outcome-equivalent to $EADA$, is the unique *constrained efficient* mechanism that gives students incentives to consent for their priorities to be violated (Dur et al., 2019).

The third approach (pursued in this paper as well) is to minimize the amount of priority violations within a class of mechanisms. This approach requires a notion of *fairness* comparison between two matchings. Consider a priority violation. This involves two students i, j and a school s, such that i has a higher priority at s, i prefers s to their assigned school, and j is assigned to s. The pair (i, j) and triplet (i, j, s) are called *justified-envy pair* and *justified-envy triplet*, respectively. Given a set of mechanisms, a mechanism in this set is *p-justified-envy minimal* (*t-justified-envy minimal*), if, for each school choice problem, the set of *justified-envy pairs* (*justified-envy triplets*) it generates is (weakly) included in the set of *justified-envy pairs* (*justified-envy triplets*) generated by any other mechanism in the same set. The *EADA* is *p-justified-envy minimal* and *t-justified-envy minimal* in the set of *Pareto efficient* mechanisms (Kwon and Shorrer, 2020).

When objects have unit capacities, the well-known *Top Trading Cycles mechanism* (*TTC*) is *p-justified-envy minimal* in the class of *efficient* and *strategy-proof* mechanisms (Abdulkadiroğlu et al., 2020). Also, this characterization of the *TTC* holds for a class of fairness comparison notions (Doğan and Ehlers, 2021b). Another notion of *fairness* comparison is based on the number of all student-school pairs in priority violations: a matching is *cardinally more fair* than another matching if the number of such student-school pairs in the former matching is less than the number of such student-school pairs in the latter matching.[2] It turns out that there is no welfare improvement over the *DA* that is *cardinally minimally unfair* among *efficient* matchings (Doğan and Ehlers, 2021a). For a matching problem with coarse priorities, finding an *efficient* matching that minimizes priority violations is NP-hard (Abdulkadiroğlu and Grigoryan, 2021).

2 The Model

We consider the standard school choice model, where the agents are students and the objects are schools (Abdulkadiroğlu and Sönmez, 2003). A school choice problem is composed of the following elements:

[2] Abraham et al. (2005) introduce the counterpart of this comparison notion for the *roommates problem.*

- a finite set of students, I,
- a finite set of schools, S,
- a capacity profile, $q = (q_s)_{s \in S}$, where q_s is the capacity of school s,
- a priority profile, $\succ = (\succ_s)_{s \in S}$, where $\succ_s$ is the strict priority order of school s over students,
- a preference profile, $P = (P_i)_{i \in I}$, where P_i is the strict preference order of student i over S.

The set of schools S includes the null-school, which represents the option of being unassigned, and is denoted by s_0. The null-school is not scarce, that is, $q_{s_0} = |I|$. Let R_i be "*the at-least-as-good-as*" relation associated with P_i, that is, $s \; R_i \; s'$ if and only if $s \; P_i \; s'$ or $s = s'$. Let $R = (R_i)_{i \in I}$. For the sake of exposition, we use $(\succ, P)$ to denote a **(school choice) problem**.

A **matching** $\mu : I \to S$ is a function such that for each $s \in S$, $|\mu^{-1}(s)| \leq q_s$. With slight abuse of notation, we use μ_i and μ_s instead of $\mu(i)$ and $\mu^{-1}(s)$, respectively.

Let $(\succ, P)$ be a problem. Matching μ is **individually rational** if, for each $i \in I$, $\mu_i \; R_i \; s_0$. Matching μ is **non-wasteful** if there does not exist a student-school pair (i, s) such that $s \; P_i \; \mu_i$ and $|\mu_s| < q_s$. Matching μ **Pareto dominates** another matching ν if, for each $i \in I$, $\mu_i \; R_i \; \nu_i$ and for some $j \in I$, $\mu_j \; P_j \; \nu_j$. Matching μ is **Pareto efficient** if it is not *Pareto dominated* by another matching. A student i has **justified-envy** under μ if there exists $s \in S$ such that $s \; P_i \; \mu_i$ and $i \; \succ_s \; j$ for some $j \in \mu_s$. Matching μ is **fair** if there does not exist a student i who has *justified-envy*. Matching μ is **stable** if it is *individually rational*, *non-wasteful*, and *fair*.

A **mechanism** is a procedure that selects a matching for each problem $(\succ, P)$. We denote the matching selected by mechanism ϕ for problem $(\succ, P)$ by $\phi(\succ, P)$. A mechanism ϕ is **stable** (**Pareto efficient**) if, for each $(\succ, P)$, $\phi(\succ, P)$ is *stable* (*Pareto efficient*). A mechanism ψ **Pareto dominates** another mechanism ϕ if, for each problem $(\succ, P)$, $\psi(\succ, P)$ *Pareto dominates* $\phi(\succ, P)$ or $\psi(\succ, P) = \phi(\succ, P)$, where the former holds for some problem.

The celebrated Deferred Acceptance (DA)[3] mechanism (Gale and Shapley, 1962) is *stable*, but not *Pareto efficient*. Indeed, there is no mechanism that is *fair* and *Pareto efficient* (Balinski and Sönmez, 1999). The outcome of the DA *Pareto dominates* any other *fair* matching.

3 Incentives for Truthful Revelation

A mechanism ϕ is **strategy-proof** if there does not exist a problem $(\succ, P)$, a student i and a preference order P'_i such that $\phi_i(\succ, P'_i, P_{-i})\ P_i\ \phi_i(\succ, P)$ where $P_{-i} = (P_j)_{j \neq i}$.

The DA is *strategy-proof* (Dubins and Freedman, 1981; Roth, 1982). In fact, the DA is the unique *strategy-proof* and *stable* mechanism. Moreover, any mechanism *Pareto dominating* the DA is not *strategy-proof* (Kesten and Kurino, 2019; Abdulkadiroğlu et al., 2009; Alva and Manjunath, 2019). Since our main focus is the class of mechanisms *Pareto dominating* the DA, we study weaker non-manipulability properties. To this end, we decompose *strategy-proofness* into two weaker notions based on possible ways to manipulate a mechanism.[4] These two notions address preference manipulation involving, separately, only the schools better and only the schools worse than the student's assignment under a mechanism. Before introducing these properties, we define two notions that we refer to throughout our analysis.

For a given preference order P_i and a school s, let $\mathcal{U}_s(P_i)$ be the set of schools ranked above s under P_i, that is, $\mathcal{U}_s(P_i) = \{s' \in S \ : \ s'\ P_i\ s\}$. Similarly, let $\mathcal{L}_s(P_i)$ be the set of schools ranked below s under P_i, that is, $\mathcal{L}_s(P_i) = \{s' \in S \ : \ s\ P_i\ s'\}$. Clearly, for each $s \in S$ and P_i, $\mathcal{U}_s(P_i) \cup \mathcal{L}_s(P_i) \cup \{s\} = S$, and $\mathcal{U}_s(P_i) \cap \mathcal{L}_s(P_i) = \emptyset$

A preference order P'_i is an **upper-manipulation** of P_i at school s, if the following are true:

a. $\mathcal{L}_s(P_i) \subseteq \mathcal{L}_s(P'_i)$, and

b. for each $s', s'' \in \mathcal{L}_s(P_i)$, $s'\ P_i\ s''$ implies $s'\ P'_i\ s''$.

[3] We provide a description of the DA in Appendix A.

[4] Recent works analyze other possible notions of decomposition of *strategy-proofness* in probabilistic settings (Bogomolnaia and Moulin, 2001; Hashimoto et al., 2014; Mennle and Seuken, 2021).

That is, if a preference order P'_i is an *upper-manipulation* of P_i at school s, then each school ranked below s under P_i is also ranked below s under P'_i and the relative ranking of those schools is the same under P'_i and P_i. A mechanism ϕ is **upper-manipulation-proof** if, for each $(\succ, P)$ and i,

$$\phi_i(\succ, P)\ R_i\ \phi_i(\succ, P'_i, P_{-i}),$$

where P'_i is an arbitrary *upper-manipulation* of P_i at school $\phi_i(\succ, P)$.

A preference order P'_i is a **lower-manipulation** of P_i at school s, if the following are true:

a. $\mathcal{U}_s(P_i) \subseteq \mathcal{U}_s(P'_i)$, and
b. for each $s', s'' \in \mathcal{U}_s(P_i)$, $s'\ P_i\ s''$ implies $s'\ P'_i\ s''$.

That is, if a preference order P'_i is a *lower-manipulation* of P_i at school s, then each school ranked above s under P_i is also ranked above s under P'_i and the relative ranking of those schools is the same under P'_i and P_i. A mechanism ϕ is **lower-manipulation-proof** if, for each $(\succ, P)$ and i,

$$\phi_i(\succ, P)\ R_i\ \phi_i(\succ, P'_i, P_{-i}),$$

where P'_i is an arbitrary *lower-manipulation* of P_i at school $\phi_i(\succ, P)$.

As it turns out, *strategy-proofness* can be decomposed into *upper-manipulation-proofness* and *lower-manipulation-proofness.*

Theorem 1. *A mechanism is strategy-proof if and only if it is upper-manipulation-proof and lower-manipulation-proof.*

We search for the set of mechanisms that *Pareto dominate* the *DA*. Since any mechanism *Pareto dominating* the *DA* fails to be *strategy-proof*, Theorem 1 implies that any such mechanism cannot be *upper-manipulation-proof* and *lower-manipulation-proof* at the same time. Our next result shows that the search for such a mechanism must involve giving up *lower-manipulation-proofness.*

Proposition 1. *There does not exist a mechanism that Pareto dominates the DA and is lower-manipulation-proof.*

By Proposition 1, a *Pareto improvement* over *stability* can only be attained at the cost of *lower-manipulation-proofness*. Given the decomposition in Theorem 1, a natural direction is to look for the *upper-manipulation-proof* mechanisms that *Pareto dominate* the DA.

4 Fairness

The core *fairness* notion in school choice is *justified-envy-freeness* (see Section 2). Since the DA *Pareto dominates* any *justified-envy-free* mechanism, any mechanism *Pareto dominating* the DA is not *justified-envy-free*. For the latter class of mechanisms, we need a weaker *fairness* axiom. To this end, we present a natural notion of comparing mechanisms in terms of *fairness*.

4.1 A Notion of Fairness Comparison

For each problem $(\succ, P)$, let $JE^{\mu}(\succ, P)$ denote the set of students who have *justified-envy* under matching μ. A mechanism ϕ has **weakly less justified-envy**[5] than another mechanism ϕ' if for each problem $(\succ, P)$,

$$JE^{\phi(\succ,P)}(\succ, P) \subseteq JE^{\phi'(\succ,P)}(\succ, P).$$

That is, if, for each problem, the set of students with *justified-envy* under mechanism ϕ is included in the set of students with *justified-envy* under mechanism ϕ', then ϕ has *weakly less justified-envy* than ϕ'.

Next, we strengthen this comparison by requiring the set inclusion relation to be strict for at least one problem. A mechanism ϕ has **less justified-envy** than ϕ' if it has *weakly less*

[5] To the best of our knowledge, this notion of comparison is first introduced by Doğan and Ehlers (2021a). By referring to this notion, they provide negative results for the $EADA$.

justified-envy than ϕ', and there exists a problem $(\succ', P')$ such that

$$JE^{\phi(\succ',P')}(\succ', P') \subset JE^{\phi'(\succ',P')}(\succ', P').$$

We are now ready to provide our notion of *fairness* comparison. Given a set of mechanisms Φ, a mechanism $\phi \in \Phi$ is **(strongly) minimally unfair in** Φ if, there is no other mechanism $\phi' \in \Phi$ that has *(weakly) less justified-envy* than ϕ.

4.2 Alternative Notions of Fairness Comparison

There are other *fairness* comparison notions used to formulate minimization of priority violations. The closest to ours is the comparison of matchings based on the sets of student pairs (and triplets of two students and the school) involved in priority violations (Abdulkadiroğlu et al., 2020; Kwon and Shorrer, 2020).

Let $(\succ, P)$ be a problem and μ a matching. If $\mu_j \; P_i \; \mu_i$ and $i \succ_{\mu_j} j$, then the triplet (i, j, μ_j) and the pair (i, j) are called **justified-envy triplet** and **justified-envy pair**, respectively. Let $JET^{\mu}(\succ, P)$ and $JEP^{\mu}(\succ, P)$ denote the sets of *justified-envy triplets* and *justified-envy pairs*, respectively.

A mechanism ϕ has **weakly less t-justified-envy** [**weakly less p-justified-envy**] than ϕ' if for each problem $(\succ, P)$, $JET^{\phi(\succ,P)}(\succ, P) \subseteq JET^{\phi'(\succ,P)}(\succ, P)$ [$JEP^{\phi(\succ,P)}(\succ, P) \subseteq JEP^{\phi'(\succ,P)}(\succ, P)$]. A mechanism ϕ has **less t-justified-envy** [**less p-justified-envy**] than ϕ' if it is *weakly less t-justified-envy* [*weakly less p-justified-envy*] than ϕ', and there exists a problem $(\succ', P')$ such that $JET^{\phi(\succ',P')}(\succ', P') \subset JET^{\phi'(\succ',P')}(\succ', P')$ [$JEP^{\phi(\succ',P')}(\succ', P') \subset JEP^{\phi'(\succ',P')}(\succ', P')$].

Given a set of mechanisms Φ, a mechanism $\phi \in \Phi$ is **(strongly) t-justified-envy minimal in** Φ if there is no other mechanism $\phi' \in \Phi$ that has *(weakly) less t-justified-envy* than ϕ. Similarly, a mechanism $\phi \in \Phi$ is **(strongly) p-justified-envy minimal in** Φ if there is no other mechanism $\phi' \in \Phi$ that has *(weakly) less p-justified-envy* than ϕ.

Next, we show the relations between the notions of *fairness* comparison.

Proposition 2.

i. Strong minimal unfairness implies strong t-justified-envy minimality and strong p-justified-envy minimality. But, neither strong t-justified-envy minimality nor p-justified-envy minimality imply minimal unfairness.

ii. Minimal unfairness, t-justified-envy minimality, and p-justified-envy minimality are independent with each other.

There are other studies conducting *fairness* comparisons and propose similar *justified-envy minimality* notions as above as well (Abdulkadiroğlu et al., 2020; Tang and Zhang, 2021). Doğan and Ehlers (2021b)'s axiomatic approach covers a class of comparison notions, including *t-justified-envy minimality* and *p-justified-envy minimality* as well as *minimal unfairness*. A different notion is *fairness* comparison based on the number of all student-school pairs in priority violations: a matching is *cardinally more fair* than another matching if the number of such student-school pairs in the former matching is less than the number of such student-school pairs in the latter matching (Doğan and Ehlers, 2021a).

5 Truthfulness, Efficiency and Fairness

The overarching theme of this work is the investigation of the "minimally unfair" among all *Pareto efficient* and *upper-manipulation-proof* mechanisms. Our main result identifies a particular mechanism, the *Efficiency Adjusted Deferred Acceptance* ($EADA$) mechanism (Kesten, 2010), as a *strongly minimally unfair* mechanism among all *Pareto efficient* and *upper-manipulation-proof* mechanisms. We first define the $EADA$.

5.1 The Efficiency Adjusted DA (EADA) Mechanism

Before defining the $EADA$, we provide some notions used in the definition. For each problem $(\succ, P)$, a student-school pair (i, s) is an **interrupting pair** if, when the DA is applied to $(\succ, P)$, student i causes another student j to be rejected at some step, say Step k, and then she is rejected from school s at some weakly later step, say Step $k' \geq k$. Given an *interrupting pair* (i, s), student i is called an **interrupter**. For each problem $(\succ, P)$, the $EADA$ selects its outcome as follows:[6]

The EADA:

Step 0: Apply the DA algorithm to problem $(\succ, P)$.

Step $k \geq 1$: Find the last step (of the DA algorithm run in Step $k-1$) in which an *interrupter* is rejected from the school for which she is an *interrupter*. Identify all *interrupting pairs* of that step. If there is no *interrupting pair*, then stop. For each identified *interrupting pair* (i, s), update the preferences of student i by removing s without changing the relative order of the remaining schools. Rerun the DA for the updated preference profile.

As both the students and schools are finite, the algorithm terminates at some step. The DA outcome in the terminal step defines the $EADA$ matching. Let $E(\succ, P)$ denote the outcome of the $EADA$ for problem $(\succ, P)$.

Remark 1. *We consider the version of the $EADA$ where each student consents to all priority violations. Thus, by Theorem 1 of Kesten (2010), the $EADA$ is Pareto efficient.*

By design, the $EADA$ *Pareto dominates* the DA.[7] By Proposition 1, this implies that the $EADA$ is not *lower-manipulation-proof* (thus, it is not *strategy-proof*). On the other hand, the $EADA$ is *upper-manipulation-proof*.

Proposition 3. *The $EADA$ is upper-manipulation-proof.*[8]

[6] In our proofs, we also utilize alternative definitions of the $EADA$ (Tang and Yu, 2014; Dur et al., 2019). These alternative definitions are provided in Appendix B.

[7] This is most easily seen through an outcome equivalent mechanism of the $EADA$, the *Top Priority (TP)* mechanism introduced by Dur et al. (2019). See Appendix B.

[8] An even stronger version of this result holds: a student cannot change the outcome of the $EADA$ via an *upper-manipulation* (see the proof of Proposition 3 in Appendix C).

By Remark 1 and Proposition 3, we have an *upper-manipulation-proof* and *Pareto efficient* mechanism, the *EADA*. The *EADA* is not necessarily the only such mechanism. But, it is special among all such mechanisms, as we demonstrate next.

5.2 The Main Result

The *EADA* is both *t-justified-envy minimal* and *p-justified-envy minimal* in the class of *Pareto efficient* mechanisms (Theorem 1 of Kwon and Shorrer (2020)). A natural question to ask is whether this result holds for stronger versions of these notions. The answer turns out to be affirmative. We provide this stronger result (Proposition 5) in Appendix D.

On the other hand, the following result shows that these properties of the *EADA* do not extend to our notion of *fairness* comparison: The *EADA* is not *minimally unfair* in the class of *Pareto efficient* mechanisms. Indeed, it is not even *minimally unfair* in a subset of *Pareto efficient* mechanisms consisting of those *Pareto dominating* the *DA*.[9]

Proposition 4. *The EADA is not (strongly) minimally unfair in the class of Pareto efficient mechanisms that Pareto dominate the DA.*

We are now ready to state our main theorem.

Theorem 2. *The EADA is strongly minimally unfair in the class of upper-manipulation-proof and Pareto efficient mechanisms.*

Proposition 4 and Theorem 2 imply that *upper-manipulation-proofness* is necessary for the *EADA* to be *(strongly) minimally unfair* in the class of *Pareto efficient* mechanisms. Moreover, an important insight of Theorem 2 is that, among all *Pareto efficient* mechanisms that *Pareto dominate* the *DA*, the *EADA* is one of the mechanisms with the least possible departure both from *fairness* (with respect to our notion of *fairnes* comparison) and from *strategy-proofness*.

[9]This result is identical to Proposition 3.(ii) of Doğan and Ehlers (2021a). We present it here for the completeness of the analysis.

5.3 Discussion

Theorem 2 and Proposition 1 can be visually summarized as in Figure 1 below.

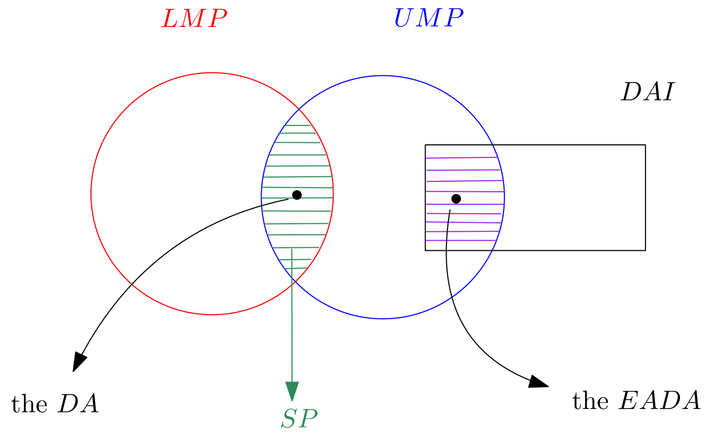

Figure 1: Visual representation of Theorem 2 and Proposition 1.

In Figure 1, UMP (LMP) is the set of *upper-manipulation-proof* (*lower-manipulation-proof*) mechanisms. By Theorem 1, the intersection (the green region) is the set of *strategy-proof* mechanisms. DAI is the set of mechanisms that *Pareto dominate* the DA. By Proposition 1, this set is disjoint with LMP. Since the $EADA$ is *upper-manipulation-proof* (Proposition 3), DAI has a non-empty intersection with UMP.

The mechanisms in the intersection of UMP and DAI (the purple region) *Pareto dominate* the DA, and they achieve strategic robustness as much as possible while doing so. By Theorem 2, the $EADA$ is on the frontier of the mechanisms in purple region in terms of *fairness*.

Why is upper-manipulation-proofness relevant? A distinguishing feature of the $EADA$ is strategic robustness: it satisfies the best incentive property one can hope for while improving over the DA; that is, it is *upper-manipulation-proof*.[10] Apart from the justification offered by Theorem 1 and Proposition 1, the value of *upper-manipulation-proofness* can be defended on various grounds.

[10]The incentive properties of the $EADA$ have not been extensively studied in the literature. A notable exception is Section V.B. of Kesten (2010), which shows that truth-telling is an equilibrium of preference revelation game induced by the $EADA$ in a "symmetric incomplete information" setting.

It can be argued that students are heuristically more inclined to strategize through their higher-ranked schools, and thus *upper-manipulations* are the more salient form of manipulations. The reason is that, as long as students' preferences are correlated, the higher-ranked schools tend to be *more popular* schools, and competition for such schools nudge students towards under-ranking them in submitted preferences. The suggestive evidence for this intuition is provided in experimental studies (Chen and Sönmez, 2006; Cerrone et al., 2022). A significant part of attempted manipulations under the DA involve *similar preference bias*, which is a type of manipulation where students under-rank schools with highest payoffs "due to the belief that participants have similar preferences, therefore, my most desirable school might also be desired by others" (Chen and Sönmez, 2006, p.218). An *upper-manipulation-proof* mechanism ensures that such deviations are not beneficial, and allows the policy-makers to advise against such attempts. Another observation points out a higher rate of truth-telling under the $EADA$ compared to the DA (Cerrone et al., 2022). The authors explain this surprising finding by arguing that it may not be easy to realize how to manipulate the $EADA$. This suggests that *lower-manipulations* are hard to detect. Thus, it is rather critical to achieve *upper-manipulation-proofness*.

A further potential benefit of *upper-manipulation-proof* mechanisms is that they provide reliable information to policy-makers. The reported rankings of students provide guidance for the school districts about the popularity and effectiveness of different programs. In particular, a school district may favor highly-demanded programs while investing on schools. Such a district can rely on the reported rankings to identify popular programs under an *upper-manipulation-proof* mechanism.

Why is our fairness comparison notion relevant? Our fairness comparison notion is based on the set of students whose priorities are violated. (While this is also referred in other works (Doğan and Ehlers, 2021a,b), to the best of our knowledge, no positive result has been provided under this notion for many-to-one matching problems.[11]) This notion is relevant

[11]On the other hand, there are some positive results for the $EADA$ in many-to-one matching problems using different comparison notions based on *justified-envy pairs* and *justified-envy triplets* (Kwon and Shorrer, 2020; Tang and Zhang, 2021). See Section 4.2 for these notions.

whenever the criterion that matters is whether or not a student's priority is violated. Given that, for a student, a single violation is sufficient to file a complaint, school districts might be concerned only with the set of such students, rather than all instances of priority violations. For these cases, our comparison notion captures school districts' approach to *fairness*.

6 Conclusion

When comparing student assignment mechanisms, economists mainly focus on three properties: fairness, efficiency and strategy-proofness. No mechanism satisfies all these three properties. In their seminal paper, Abdulkadiroğlu and Sönmez (2003) suggested school districts adopt the Deferred Acceptance (DA) if they value fairness over efficiency and adopt the Top Trading Cycles (TTC) if they value efficiency over fairness. Findings in this paper provides a rationale to consider another mechanism: the $EADA$ (Kesten, 2010). The $EADA$ is in the frontier in terms of all these three properties. Any mechanism outperforming the $EADA$ based on any of these properties performs poorer based on another one.

Appendix A The Deferred Acceptance (DA) Mechanism

Given a problem $(\succ, P)$, the DA selects its outcome as follows:

The Deferred Acceptance:

Step 1. Each student applies to their best school. Each school tentatively accepts the top priority applicants up to its quota and rejects the rest.

In general,

Step k. Each rejected student applies to their next best school. Among the current step applicants and the tentatively accepted ones, each school tentatively accepts the top priority applicants up to its quota and rejects the rest.

The algorithm terminates whenever each student is assigned to a school. The matching in the terminal step constitutes the DA outcome.

Appendix B Alternative Definitions of the EADA

There is an alternative definition of the $EADA$, the **simplified** $EADA$ (Tang and Yu, 2014), which is arguably more tractable than the $EADA$ (Kesten, 2010).

Given a problem $(\succ, P)$ and a matching μ, a school s is **undemanded** if there does not exist a student i such that $s \; P_i \; \mu_i$. Given a problem $(\succ, P)$, the *simplified* $EADA$ selects its outcome as follows:

The simplified EADA:

Step 0: Apply the DA to problem $(\succ, P)$. If there is no *undemanded* school, then the algorithm terminates. Otherwise, remove an *undemanded* school along with the assignees.

Step $k > 1$: Apply the DA to the reduced problem obtained in Step $k-1$. If there is no *undemanded* school, then the algorithm terminates. Otherwise, remove an *undemanded* school along with the assignees.

This is a slightly modified version of the *simplified EADA* by Tang and Yu (2014) such that, instead of removing all *undemanded* schools with their assignees altogether, they are removed one by one. This modification does not alter the outcome.

Since the numbers of students and schools are finite, the algorithm terminates in finitely many steps. The DA outcome in its terminal step along with the previously removed students with their matched schools constitute the algorithm's outcome.

There is also another class of mechanisms, where the outcome of each member *Pareto dominates* the outcome of the DA by violating only allowable priorities, which are exogenously defined (Dur et al., 2019). There exists a mechanism in this class, called the **Top Priority mechanism** (TP), which is outcome equivalent to the $EADA$ (Dur et al., 2019).

Given a problem $(\succ, P)$ and a matching μ, let G be the graph with the set of vertices I such that i points to j if and only if $\mu_j \ P_i \ \mu_i$. A school s is **underdemanded** if either s is *undemanded* or no student on a path to a student in μ_s is part of a cycle in G. Given a problem $(\succ, P)$, the TP selects its outcome as follows:

The Top Priority mechanism (TP):

Step 0: Apply the DA to problem $(\succ, P)$. Let $\mu_0 = DA(\succ, P)$.

Step $k > 0$: Let U^{k-1} be the set of students assigned to *underdemanded* schools under μ^{k-1}. Let $D_s^{k-1} = \{i \in I \setminus U^{k-1} \ : \ s \ P_i \ \mu_i^{k-1}\}$. Let G_{TP}^{k-1} be the graph such that $i \in D_s^{k-1}$ points to $j \in \mu_s^{k-1}$ if and only if i has the highest priority for s among the students in D_s^{k-1}. A cycle in G_{TP}^{k-1} is a **top priority cycle**. If there does not exist such a cycle, then the mechanism terminates. Otherwise, we implement a *top priority cycle*, and assign each student in that cycle to the school of the student she points to. Let μ^k be the matching obtained from μ^{k-1} by implementing a cycle.

The outcome of the TP does not depend on the order of cycles implemented, and thus it is a well-defined mechanism (Dur et al., 2019).

Appendix C Proofs

Proof of Theorem 1. Let ϕ be a mechanism. If ϕ is *strategy-proof*, then by definition it is *upper-manipulation-proof* and *lower-manipulation-proof*.

To prove the other direction, suppose, towards a contradiction, that ϕ is *upper-manipulation-proof* and *lower-manipulation-proof*, but not *strategy-proof*. Then, there exists a problem $(\succ, P)$, a student i and a preference order P_i' such that $\phi_i(\succ, P_i', P_{-i})\ P_i\ \phi_i(\succ, P)$. Let $s' \equiv \phi_i(\succ, P_i', P_{-i})$ and $s \equiv \phi_i(\succ, P)$. We define

$$\tilde{S} \equiv \mathcal{U}_s(P_i') \cap \mathcal{L}_s(P_i) = \{\tilde{s} \in S : \tilde{s}\ P_i'\ s\ P_i\ \tilde{s}\}.$$

Claim 1. *$\tilde{S}$ is non-empty.*

Proof. Suppose, towards a contradiction, that $\tilde{S} = \emptyset$. This implies that for each $\tilde{s} \in \mathcal{U}_s(P_i')$, $\tilde{s} \notin \mathcal{L}_s(P_i)$. Thus,

$$\mathcal{U}_s(P_i') \subseteq \mathcal{U}_s(P_i).$$

Similarly,

$$\mathcal{L}_s(P_i) \subseteq \mathcal{L}_s(P_i').$$

Let P_i'' be the preference order such that:

i. $\mathcal{L}_s(P_i'') = \mathcal{L}_s(P_i)$ and $\mathcal{U}_s(P_i'') = \mathcal{U}_s(P_i)$,

ii. for each $s', s'' \in \mathcal{U}_s(P_i)$, $s'\ P_i\ s''$ implies $s'\ P_i''\ s''$,

iii. for each $s', s'' \in \mathcal{L}_s(P''_i)$, $s' \ P'_i \ s''$ implies $s' \ P''_i \ s''$.

By (i) and (ii), P''_i is a *lower-manipulation* of P_i at s. Let $s'' \equiv \phi_i(\succ, P''_i, P_{-i})$. Since ϕ is *lower-manipulation-proof*, $s \ R_i \ s''$, which implies $s'' \in \mathcal{L}_s(P_i) \cup \{s\}$. Moreover, since $s' \ \ P_i \ \ s$, $s' \in \mathcal{U}_s(P_i)$. By (i), these observations imply $s'' \in \mathcal{L}_s(P''_i) \cup \{s\}$ and $s' \in \mathcal{U}_s(P''_i)$. Thus, $s' \ \ P''_i \ \ s''$.

Finally, since $s'' \in \mathcal{L}_s(P_i) \cup \{s\}$, by (i) and (iii), P'_i is an *upper-manipulation* of P''_i at s''. Since ϕ is *upper-manipulation-proof*, this implies that $s'' \ \ R''_i \ s'$, which is a contradiction. We conclude that $\tilde{S} \neq \emptyset$. □

Let P^1_i be the preference order, constructed based on P_i, P'_i and $\tilde{S}$, such that:

I. $\mathcal{U}_s(P^1_i) = \mathcal{U}_s(P_i) \cup \tilde{S}$ and $\mathcal{L}_s(P^1_i) = \mathcal{L}_s(P_i) \setminus \tilde{S}$,

II. for each $s' \in \mathcal{U}_s(P_i)$ and $s'' \in \tilde{S}$, $s' \ P^1_i \ s''$,

III. for each $s', s'' \in \mathcal{U}_s(P_i)$, $s' \ P_i \ s''$ implies $s' \ P^1_i \ s''$,

IV. for each $s', s'' \in \tilde{S}$, $s' \ P'_i \ s''$ implies $s' \ P^1_i \ s''$,

V. for each $s', s'' \in \mathcal{L}_s(P_i) \setminus \tilde{S}$, $s' \ P_i \ s''$ implies $s' \ P^1_i \ s''$.

By (I) and (III), P^1_i is a *lower-manipulation* of P_i at s. Let $s^1 \equiv \phi_i(\succ, P^1_i, P_{-i})$. Since ϕ is *lower-manipulation-proof*, $s \ \ R_i \ \ s^1$. Also recall that $s' \in \mathcal{U}_s(P_i)$. This, combined with (II), implies that $s' \ \ P^1_i \ \ s^1$. Thus, for problem $(\succ, P^1_i, P_{-i})$, the preference order P'_i is a profitable deviation for student i. Based on this observation, we define

$$\tilde{S}^1 \equiv \mathcal{U}_{s^1}(P'_i) \cap \mathcal{L}_{s^1}(P^1_i) = \{\tilde{s} \in S : \tilde{s} \ P'_i \ s^1 \ P^1_i \ \tilde{s}\}$$

Since ϕ is both *upper-manipulation-proof* and *lower-manipulation-proof*, and also P'_i is a profitable deviation for i in problem $(\succ, P^1_i, P_{-i})$, Claim 1 applies to $\tilde{S}^1$ as well. Thus, $\tilde{S}^1 \neq \emptyset$.

Our next claim is that $s^1 \notin \tilde{S}$. We prove this claim by showing that $s^1 \in \tilde{S}$ contradicts to $\tilde{S}^1 \neq \emptyset$. Suppose, towards a contradiction, that $s^1 \in \tilde{S} = \mathcal{U}_s(P_i') \cap \mathcal{L}_s(P_i)$. This implies:

$$s^1 \; P_i' \; s. \tag{1}$$

Moreover, we know that $\tilde{S}^1 \neq \emptyset$. Let $\tilde{s}^1 \in \tilde{S}^1$. By definition of the set $\tilde{S}^1$,

$$\tilde{s}^1 \; P_i' \; s^1, \tag{2}$$

$$s^1 \; P_i^1 \; \tilde{s}^1. \tag{3}$$

First, note that (1) and (2) imply $\tilde{s}^1 \; P_i' \; s$. Thus, $\tilde{s}^1 \neq s$. Second, since $s^1 \in \tilde{S}$, by (I), $s^1 \;\; P_i^1 \;\; s$. We next investigate (3) under two possible cases:

Case 1: $s \; P_i^1 \; \tilde{s}^1$.
First, since, by (I), $\mathcal{L}_s(P_i^1) = \mathcal{L}_s(P_i) \setminus \tilde{S}$, we have $\mathcal{L}_s(P_i^1) \subset \mathcal{L}_s(P_i)$. Thus, $\tilde{s}^1 \in \mathcal{L}_s(P_i^1)$ implies $\tilde{s}^1 \in \mathcal{L}_s(P_i)$. Second, by (1) and (2), $\tilde{s}^1 \;\; P_i' \;\; s$. Thus, $\tilde{s}^1 \in \mathcal{U}_s(P_i')$. By combining these observations, we conclude that $\tilde{s}^1 \in \tilde{S}$.

Case 2: $\tilde{s}^1 \; P_i^1 \; s$.
By (3), $s^1 \; P_i^1 \; \tilde{s}^1 \; P_i^1 \; s$. Since $s^1 \in \tilde{S}$, by (I) and (II), $\tilde{s}^1 \in \tilde{S}$.

In either case, $\tilde{s}^1 \in \tilde{S}$. Then, both $s^1 \in \tilde{S}$ and $\tilde{s}^1 \in \tilde{S}$. By (IV) and (2), $\tilde{s}^1 \; P_i^1 \; s^1$. But, this contradicts to (3).

We conclude that $s^1 \notin \tilde{S}$. Moreover, $s \; R_i \; s^1$. By (I), these imply $s^1 \in \mathcal{L}_s(P_i^1) \cup \{s\}$. Thus,

$$\mathcal{L}_{s^1}(P_i^1) \subseteq \mathcal{L}_s(P_i^1). \tag{4}$$

Moreover, since $\tilde{S} \neq \emptyset$, by (I),

$$\mathcal{L}_s(P_i^1) \subset \mathcal{L}_s(P_i). \tag{5}$$

By combining (4) and (5), we obtain

$$\mathcal{L}_{s^1}(P_i^1) \subset \mathcal{L}_s(P_i).$$

We repeat the argument to obtain a sequence of sets: For each $k \geq 2$, we

1. construct P_i^k (based on P_i^{k-1}, P_i', and $\tilde{S}^{k-1}$) and $\tilde{S}^k$, the same way we constructed P_i^1 (based on P_i, P_i' and $\tilde{S}$) and $\tilde{S}^1$, and

2. conclude, by Claim 1, that $\tilde{S}^k \neq \emptyset$.

Let $s^k \equiv \phi_i(\succ, P_i^k, P_{-i})$. By the same argument, $s^k \notin \tilde{S}^{k-1}$ and

$$\mathcal{L}_{s^k}(P_i^k) \subset \mathcal{L}_{s^{k-1}}(P_i^{k-1}).$$

Thus,

$$\mathcal{L}_{s^k}(P_i^k) \subset \mathcal{L}_{s^{k-1}}(P_i^{k-1}) \subset \ldots \mathcal{L}_{s^2}(P_i^2) \subset \mathcal{L}_{s^1}(P_i^1).$$

Since each inclusion above is proper and S is finite, for some $k \geq 1$, we have $\mathcal{L}_{s^k}(P_i^k) = \emptyset$. But, this implies $\tilde{S}^k \equiv \mathcal{U}_{s^k}(P_i') \cap \mathcal{L}_{s^k}(P_i^k) = \emptyset$, which is contradiction. □

Proof of Proposition 1. Suppose, towards a contradiction, that there exists a mechanism ψ which *Pareto dominates* the DA and is *lower-manipulation-proof.* Let $(\succ, P)$ be a problem where $\psi(\succ, P) \neq DA(\succ, P)$. We first show that any unassigned student at the matching $DA(\succ, P)$ cannot be better off at $\psi(\succ, P)$.

Claim 2. *Let $(\succ, P)$ be a problem. Let μ be a matching that Pareto dominates the matching $DA(\succ, P)$. If $DA_i(\succ, P) = s_0$, then $\mu_i = s_0$.*

Proof. First, note that, by *individual rationality* of the DA, no assigned student at $DA(\succ, P)$ becomes better off by being unassigned at μ. This implies that, if an unassigned student at $DA(\succ, P)$ is better off at μ, then there exists at least one school s such that $|\mu_s| > |DA_s(\succ, P)|$. Since the set μ_s contains at least one student who is better off at μ than at $DA(\succ, P)$, this contradicts to *non-wastefulness* of the DA. $\square$

Let I' be the students who are better off at $\psi(\succ, P)$. Since $\psi(\succ, P)$ *Pareto dominates* $DA(\succ, P)$, by Claim 2, I' does not include any unassigned student at $DA(\succ, P)$.

Let G be the graph with the set of vertices I' and the edges such that $i \in I'$ points to $j \in I'$ if and only if $\psi_i(\succ, P) = DA_j(\succ, P)$. For each $i \in I'$, since $\psi_i(\succ, P)\ P_i\ DA_i(\succ, P)$, by *non-wastefulness* of the DA, whenever i is better off at $\psi(\succ, P)$ than at $DA(\succ, P)$, there exists another student j such that $DA_j(\succ, P) = \psi_i(\succ, P)$ and $\psi_j(\succ, P) \neq DA_j(\succ, P)$. Thus, in the graph G, whenever i points to j, the latter points to a student as well. Thus, G contains at least one cycle. Let us call any cycle in G as an **improvement cycle**. By construction of G, the matching $\psi(\succ, P)$ is obtained by executing a sequence of these *improvement cycles*.

Let $i_1 \to i_2 \to ... \to i_n \to i_1$ be an *improvement cycle*. That is, for each $k \in \{1, \dots, n-1\}$, $\psi_{i_k}(\succ, P) = DA_{i_{k+1}}(\succ, P)$, and $\psi_{i_n}(\succ, P) = DA_{i_1}(\succ, P)$. Let $s_1 \equiv DA_{i_1}(\succ, P)$ and $s_2 \equiv DA_{i_2}(\succ, P)$. By definition of an *improvement cycle*, $s_2 \in \mathcal{U}_{s_1}(P_{i_1})$. Moreover, since $i_1 \in I'$, by Claim 2, $s_1 \neq s_0$. Also, since the DA is *individually rational*, $s_0 \in \mathcal{L}_{s_1}(P_{i_1})$.

Let P'_{i_1} be the preference order such that:

i. $\mathcal{U}_{s_0}(P'_{i_1}) = \mathcal{U}_{s_1}(P_{i_1})$ and $\mathcal{L}_{s_0}(P'_i) = (\mathcal{L}_{s_1}(P_{i_1}) \cup \{s_1\}) \setminus \{s_0\}$,

ii. for each $s', s'' \in \mathcal{U}_{s_0}(P'_{i_1})$, $s'\ P_{i_1}\ s''$ implies $s'\ P'_{i_1}\ s''$.

Let $P' = (P'_{i_1}, P_{-i_1})$. Since the DA is *individually rational*, $DA_{i_1}(\succ, P')\ R'_{i_1}\ s_0$, which is equivalent to $DA_{i_1}(\succ, P') \in \mathcal{U}_{s_0}(P'_{i_1}) \cup \{s_0\}$. By (i), this implies $DA_{i_1}(\succ, P') \in \mathcal{U}_{s_1}(P_{i_1}) \cup \{s_0\}$. If $DA_{i_1}(\succ, P') \in \mathcal{U}_{s_1}(P_{i_1})$, then for problem $(\succ, P)$, i_1 is better off by revealing P'_{i_1} rather than their true preference order P_{i_1}. This contradicts to the DA being *strategy-proof*. Thus, $DA_{i_1}(\succ, P') = s_0$. Since ψ *Pareto dominates* the DA, $\psi_{i_1}(\succ, P')\ R_{i_1}\ DA_{i_1}(\succ, P')$. If $\psi_{i_1}(\succ, P')\ P_{i_1}\ DA_{i_1}(\succ, P')$, then i_1 is in the set of students who are better off at $\psi(\succ, P')$, contradicting Claim 2. Thus, $\psi_{i_1}(\succ, P') = DA_{i_1}(\succ, P') = s_0$.

By (i) and (ii), P_{i_1} is a *lower-manipulation* of P'_{i_1} at $\psi_{i_1}(\succ, P')$. Moreover,

$$s_2 = \psi_{i_1}(\succ, P)\ P'_{i_1}\ \psi_{i_1}(\succ, P') = s_0, \tag{6}$$

which contradicts to *lower-manipulation-proofness* of ψ. □

Proof of Proposition 2. *i*. It is immediate to see from their definitions that *strong minimal unfairness* implies both *strong t-justified-envy minimality* and *strong p-justified-envy minimality*.

For the counterpart, consider a problem where $I = \{i_1, i_2, i_3\}$, $S = \{a, b\}$, $q = (1, 1)$, and the priorities and preferences are:

P_{i_1}	P_{i_2}	P_{i_3}	$\succ_a$	$\succ_b$
a	a	b	i_1	i_1
b	s_0	s_0	i_2	i_3
s_0	b	a	i_3	i_2

Let ψ and ϕ be mechanisms such that $\psi_{i_1}(\succ, P) = s_0$, $\psi_{i_2}(\succ, P) = a$, $\psi_{i_3}(\succ, P) = b$; $\phi_{i_1}(\succ, P) = s_0$, $\phi_{i_2}(\succ, P) = a$, and $\phi_{i_3}(\succ, P) = s_0$. For each problem other than $(\succ, P)$, ψ produces a *stable* matching while ϕ produces a matching with *justified envy* whenever possible. Note that $JET^{\phi(\succ,P)}(\succ, P) = \{(i_1, i_2, a)\}$ and $JET^{\psi(\succ,P)}(\succ, P) = \{(i_1, i_2, a), (i_1, i_3, b)\}$. For the class of mechanisms $\Phi = \{\psi, \phi\}$, ϕ is both *strongly t-justified-envy minimal* and *strongly p-justified-envy minimal* in Φ. But, ψ has *less justified-envy* than ϕ. Thus, ϕ is not *minimally unfair* in Φ.

ii. The properties *t-justified-envy minimality* and *p-justified-envy minimality* are independent (Proposition 1 of Kwon and Shorrer (2020)). By Part (*i*), neither *t-justified-envy minimality* nor *p-justified-envy minimality* implies *minimal unfairness*.

We complete the proof by showing that *minimal unfairness* implies neither *t-justified-envy minimality* nor *p-justified-envy minimality*. Let $I = \{i_1, i_2, i_3\}$, $S = \{a, b, c\}$ and $q = (1, 1, 1)$.

Let ψ and ψ' be two mechanisms which select the same matchings for each problem except the following:

P_{i_1}	P_{i_2}	P_{i_3}	$\succ_a$	$\succ_b$	$\succ_c$
a	a	c	i_1	i_1	i_1
b	b	b	i_2	i_2	i_2
c	c	a	i_3	i_3	i_3

Let $\nu = \psi(\succ, P)$ and $\nu' = \psi'(\succ, P)$. Let $\nu_{i_1} = b$, $\nu_{i_2} = a$, $\nu_{i_3} = c$ and $\nu'_{i_1} = c$, $\nu'_{i_2} = a$, $\nu'_{i_3} = b$. Note that $JET^{\nu}(\succ, P) = \{(i_1, i_2, a)\}$, $JEP^{\nu}(\succ, P) = \{(i_1, i_2)\}$, $JE^{\nu}(\succ, P) = \{i_1\}$, $JET^{\nu'}(\succ, P) = \{(i_1, i_2, a), (i_1, i_3, b)\}$, $JEP^{\nu'}(\succ, P) = \{(i_1, i_2), (i_1, i_3)\}$, and $JE^{\nu'}(\succ, P) = \{i_1\}$. Thus, while both ψ and ψ' are *minimally unfair* in $\Phi = \{\psi, \psi'\}$, only ψ is both *p-justified-envy minimal*, and *t-justified-envy minimal*. Thus, *minimal unfairness* implies neither *t-justified-envy minimality* nor *p-justified-envy minimality*. □

Proof of Proposition 3. Let $(\succ, P)$ be a problem. Let i be a student, and P'_i an *upper-manipulation* of P_i at $s = E_i(\succ, P)$. Let $\hat{P}_i$ be the preference order and $\hat{P} = (\hat{P}_i, P_{-i})$ such that:

i. $\mathcal{U}_s(\hat{P}_i) = \mathcal{U}_s(P'_i)$ and $\mathcal{L}_s(\hat{P}_i) = \mathcal{L}_s(P'_i)$,

ii. for each $s', s'' \in \mathcal{L}_s(\hat{P}_i)$, $s' \; P'_i \; s''$ implies $s' \; \hat{P}_i \; s''$,

iii. for each $s', s'' \in \mathcal{U}_s(\hat{P}_i)$, $s' \; P_i \; s''$ implies $s' \; \hat{P}_i \; s''$.

Claim 3. *$E(\succ, P) = E(\succ, \hat{P})$.*

Proof. We first define **Maskin monotonicity**. A preference order P'_i is a **monotonic transformation** of P_i at $s \in S$, if each school ranked above s under P'_i is also ranked above s under P_i. A preference profile P' is a *monotonic transformation* of P at matching μ, if for each i, P'_i is a *monotonic transformation* of P_i at μ_i. Mechanism ϕ satisfies **Maskin monotonicity** if, for each problem $(\succ, P)$ and for each *monotonic transformation* P' of P

at $\phi(\succ, P)$, $\phi(\succ, P') = \phi(\succ, P)$. Mechanism ϕ satisfies **weak Maskin monotonicity** if, for each problem $(\succ, P)$ and for each *monotonic transformation* P' of P at $\phi(\succ, P)$, we have that for each i, $\phi_i(\succ, P')\ R'_i\ \phi_i(\succ, P)$.

We are now ready to start with the proof. Let $\mu = DA(\succ, P)$. Since the *EADA (weakly) Pareto dominates* the *DA*, for each $j \in I$, $E_j(\succ, P)\ R_j\ \mu_j$. Let s_k be the *undemanded* school and I_k the set of students removed at Step k of the *simplified EADA* for problem $(\succ, P)$.

The preference profile $\hat{P}$ is a *monotonic transformation* of P at school s. Since $s\ R_i\ \mu_i = DA_i(\succ, P)$, $\hat{P}$ is also a *monotonic transformation* of P at school μ_i. This implies the following:

I. Since the *DA* is *weakly Maskin monotonic* (Kojima and Manea, 2010), for each $j \in I$, $DA_j(\succ, \hat{P})\ \hat{R}_j\ \mu_j$.

II. μ is *non-wasteful* also for the problem $(\succ, \hat{P})$.[12]

III. At μ, each *undemanded* school for $(\succ, P)$ is also *undemanded* for $(\succ, \hat{P})$.

Consider the problem $(\succ, \hat{P})$. Since, by (III), s_1 is *undemanded* at μ, and by (I), $DA(\succ, \hat{P})$ *Pareto dominates* μ, there is no student who is assigned to s_1 under $DA(\succ, \hat{P})$ but not assigned to s_1 under μ. Thus,

$$DA_{s_1}(\succ, \hat{P}) \subseteq \mu_{s_1}. \tag{7}$$

Suppose there exists a student in $\mu_{s_1} \setminus DA_{s_1}(\succ, \hat{P})$. Since the inclusion in (7) holds for each *undemanded* school, there exists at least one school such that the number of students assigned to it at $DA(\succ, \hat{P})$ is strictly higher than at μ. Since μ is *non-wasteful* by (II), this is a contradiction. Thus,

$$DA_{s_1}(\succ, \hat{P}) = \mu_{s_1}. \tag{8}$$

[12] The reason is the following: First, note that, P and $\hat{P}$ coincide except for i. Also, since $\hat{P}_i$ is a *monotonic transformation* of P_i at μ_i, each school in $\mathcal{U}_{\mu_i}(\hat{P}_i)$ is also in $\mathcal{U}_{\mu_i}(P_i)$. Thus, μ is *non-wasteful* for $(\succ, P)$ implies μ is *non-wasteful* also for $(\succ, \hat{P})$.

Since, by (III), s_1 is *undemanded*, the equality in (8) implies that school s_1 and students in I_1 can be removed in the first step of the *simplified EADA* (when applied to the problem $(\succ, \hat{P})$). Thus, for each $j \in I_1$, $E_j(\succ, P) = E_j(\succ, \hat{P})$. If $i \in I_1$, we are done. Otherwise, we consider students in $I \setminus I_1$ and the reduced problem.

We then consider the preference profiles P_{-I_1} and $(\hat{P}_i, P_{-(i \cup I_1)})$. Following the same reasoning as above, if $j \in I_2$, then j is assigned to an *undemanded* school at $DA(\succ, \hat{P}_i, P_{-(i\cup I_1)})$ and $DA_j(\succ, \hat{P}_i, P_{-(i\cup I_1)}) = E_j(\succ, P)$. Thus, we can remove all students in I_2 when the *simplified EADA* is applied to problem $(\succ, \hat{P})$ before all other students. If $i \in I_2$, we are done. Otherwise, we consider students in $I \setminus (I_1 \cup I_2)$ and the reduced problem.

By continuing this process in the obtained reduced problems, we conclude that each student is assigned to the same school under $E(\succ, P)$ and $E(\succ, \hat{P})$. □

Claim 4. *Let $(\succ, P^1)$ and $(\succ, P^2)$ be two problems such that $P^1_{-i} = P^2_{-i}$. Let $\mu^1 = DA(\succ, P^1)$ and $\mu^2 = DA(\succ, P^2)$. If*

$$\mathcal{U}_{\mu^1_i}(P^1_i) = \mathcal{U}_{\mu^1_i}(P^2_i), \tag{9}$$

and

$$\mathcal{L}_{\mu^1_i}(P^1_i) = \mathcal{L}_{\mu^1_i}(P^2_i), \tag{10}$$

then $\mu^1 = \mu^2$.

Proof. Suppose, for problem $(\succ, P^1)$, a student has *justified-envy* at μ^2. By (9), the same student has *justified-envy* at μ^1 for problem $(\succ, P^1)$. This contradicts to *stability* of the DA. Thus, for problem $(\succ, P^1)$, μ^2 is *stable*.

Suppose $\mu^2_i \neq \mu^1_i$. If $\mu^2_i \in \mathcal{U}_{\mu^1_i}(P^1_i)$, then P^2_i is a profitable deviation for i in problem $(\succ, P^1)$, contradicting to *strategy-proofness* of the DA. If $\mu^2_i \in \mathcal{L}_{\mu^1_i}(P^1_i)$, then, by (10), $\mu^2_i \in \mathcal{L}_{\mu^1_i}(P^2_i)$. Then, P^1_i is a profitable deviation for i in problem $(\succ, P^2)$, again contradicting to *strategy-*

proofness of the DA. Thus,

$$\mu_i^2 = \mu_i^1. \tag{11}$$

For problem $(\succ, P^2)$, since μ^1 is *stable* and μ^2 is the SOSM, μ^2 *(weakly) Pareto dominates* μ^1. By (11), this implies that μ^2 also *Pareto dominates* μ^1 under $(\succ, P^1)$. Thus, for problem $(\succ, P^1)$, if μ^2 is *stable*, then, since μ^1 is the SOSM, $\mu^2 = \mu^1$. We next show that, for problem $(\succ, P^1)$, μ^2 is *stable*, and thus, $\mu^2 = \mu^1$.

Since μ^2 is the SOSM for problem $(\succ, P^2)$, no student has *justified envy* at μ^2. This implies, by (9) and (11), that i does not have *justified envy* at μ^2 for problem $(\succ, P^1)$. Since $P^2_{-i} = P^1_{-i}$, and no student $j \neq i$ has *justified envy* at μ^2 under $(\succ, P^2)$, no student $j \neq i$ has *justified-envy* at μ^2 under $(\succ, P^1)$ either. Thus, for problem $(\succ, P^1)$, μ^2 is *stable*, and thus, $\mu^2 = \mu^1$. □

Claim 5. *$E(\succ, P') = E(\succ, \hat{P})$.*

Proof. Let $\mu' = DA(\succ, P')$ and $\hat{\mu} = DA(\succ, \hat{P})$. By Claim 3, $E_i(\succ, \hat{P}) = E_i(\succ, P)$. Thus, $E_i(\succ, \hat{P}) = s$. Since the *EADA (weakly) Pareto dominates* the DA, for each $j \in I$, $E_j(\succ, \hat{P})\ \hat{R}_j\ \hat{\mu}_j$. Thus,

$$s\ \hat{R}_i\ \hat{\mu}_i. \tag{12}$$

By (i) and (ii) in the construction of $\hat{P}_i$, and (12), we have

$$\mathcal{U}_{\hat{\mu}_i}(\hat{P}_i) = \mathcal{U}_{\hat{\mu}_i}(P'_i), \tag{13}$$

and also

$$\mathcal{L}_{\hat{\mu}_i}(\hat{P}_i) = \mathcal{L}_{\hat{\mu}_i}(P'_i). \tag{14}$$

By (13) and (14), Claim 4 implies that $\mu' = \hat{\mu}$. Thus, in Step 0 of the TP, $\hat{\mu}$ is selected under both $(\succ, \hat{P})$ and $(\succ, P')$. Moreover, by (12), and (13), i's preference order over the schools $(\mathcal{U}_{\hat{\mu}_i}(P'_i) \cup \{\hat{\mu}_i\}) \setminus \mathcal{U}_s(P'_i)$ is the same under both $\hat{P}_i$ and P'_i. Since the preference orders of students in $I \setminus \{i\}$ remain unchanged, this observation generalizes: for each student $j \in I$, j's preference order over the schools ranked below $E_j(\succ, \hat{P})$ and above $\hat{\mu}_j$ are the same in problems $(\succ, \hat{P})$ and $(\succ, P')$. Thus, the *top priority cycles* under $(\succ, \hat{P})$ and $(\succ, P')$ are identical, and all students are assigned to the same schools, implying $E(\succ, P') = E(\succ, \hat{P})$. □

Thus, Claims 3 and 5 imply that $E(\succ, P) = E(\succ, \hat{P}) = E(\succ, P')$. Thus, no student can be better off by an *upper-manipulation*, hence the $EADA$ is *upper-manipulation-proof.* □

Proof of Proposition 4. Consider a problem where $I = \{i_1, ..., i_6\}$ and $S = \{a, b, c, d\}$, where each school except d has a unit capacity, and $q_d = 2$. Let the preferences and priorities be as follows:

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}	P_{i_6}	$\succ_a$	$\succ_b$	$\succ_c$	$\succ_d$
a	b	b	c	c	d	i_2	i_1	i_6	i_4
b	a	c	b	d	c	$\vdots$	i_3	i_3	i_3
$\vdots$	$\vdots$	d	d	$\vdots$	$\vdots$	$\vdots$	i_4	i_4	i_5
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Let ψ be a mechanism such that it produces the same outcome as the $EADA$ for each problem except $(\succ, P)$ above. Let $\psi(\succ, P) = \mu$ where $\mu_{i_1} = a$, $\mu_{i_2} = b$, $\mu_{i_3} = d$, $\mu_{i_4} = c$, $\mu_{i_5} = s_0$, and $\mu_{i_6} = d$. On the other hand, $E(\succ, P) = \mu'$ where $\mu'_{i_1} = a$, $\mu'_{i_2} = b$, $\mu'_{i_3} = c$, $\mu'_{i_4} = d$, $\mu'_{i_5} = s_0$, and $\mu'_{i_6} = d$. Both μ and μ' *Pareto dominate* the DA outcome, say μ'', where $\mu''_{i_1} = b$, $\mu''_{i_2} = a$, $\mu''_{i_3} = d$, $\mu''_{i_4} = d$, $\mu''_{i_5} = s_0$, and $\mu''_{i_6} = c$. Also, $JE^{\mu}(\succ, P) = \{i_3, i_5\}$ and $JE^{\mu'}(\succ, P) = \{i_3, i_4, i_5\}$. Thus, ψ is *Pareto efficient* and both *Pareto dominates* the DA and has *less justified-envy* than the $EADA$, showing that the $EADA$ is not *(strongly) minimally unfair* in the class of *Pareto efficient* mechanisms that *Pareto dominate* the DA. □

Proof of Theorem 2. Suppose, towards a contradiction, that there exists another *Pareto* *efficient* and *upper-manipulation-proof* mechanism ϕ that has *weakly less justified-envy* than the $EADA$. Thus, there exists a problem $(\succ, P)$ such that $E(\succ, P) \neq \phi(\succ, P)$. The $EADA$ is *Pareto efficient* and *Pareto dominates* the DA. Since ϕ is also *Pareto efficient*, $E(\succ, P) \neq \phi(\succ, P)$ implies $\phi(\succ, P) \neq DA(\succ, P)$. We consider two possible cases.

Case 1: $\phi(\succ, P)$ does not Pareto dominate $DA(\succ, P)$. There exists a student i_1 such that $DA_{i_1}(\succ, P)\ P_{i_1}\ \phi_{i_1}(\succ, P)$. Let $s_1 \equiv DA_{i_1}(\succ, P)$.

Let P'_{i_1} be the preference order, constructed based on P_{i_1} and s_1, and $P' = (P'_{i_1}, P_{-i_1})$ such that:

i. $\mathcal{U}_{s_1}(P'_{i_1}) = \emptyset$ and $\mathcal{L}_{s_1}(P'_{i_1}) = \mathcal{L}_{s_1}(P_{i_1}) \cup \mathcal{U}_{s_1}(P_{i_1})$,

ii. for each $s', s'' \in \mathcal{L}_{s_1}(P'_{i_1})$, $s'\ P_{i_1}\ s''$ implies $s'\ P'_{i_1}\ s''$.

The preference order P'_{i_1} is an *upper-manipulation* of P_{i_1} at s_1, and also an *upper-manipulation* of P_{i_1} at $\phi_{i_1}(\succ, P)$.

Since the DA is *strategy-proof*, $DA_{i_1}(\succ, P') = s_1$.[13] Since the $EADA$ *Pareto dominates* the DA and s_1 is the most preferred school under P'_{i_1}, $s_1 = E_{i_1}(\succ, P')$. Moreover, since ϕ is *upper-manipulation-proof*, $\phi_{i_1}(\succ, P)\ R_{i_1}\ \phi_{i_1}(\succ, P')$. Also, since $s_1\ P_{i_1}\ \phi_{i_1}(\succ, P)$, this implies $s_1\ P_{i_1}\ \phi_{i_1}(\succ, P')$. Then, by (i) and (ii), $s_1\ P'_{i_1}\ \phi_{i_1}(\succ, P')$. Thus,

$$s_1 = E_{i_1}(\succ, P') = DA_{i_1}(\succ, P')\ P'_{i_1}\ \phi_{i_1}(\succ, P').$$

If there exists j such that $i_1 \succ_{s_1} j$ and $\phi_j(\succ, P') = s_1$, then i_1 has *justified-envy* at $\phi(\succ, P')$ but not at $E(\succ, P')$. This is a contradiction to ϕ having *weakly less justified-envy* than the $EADA$, and the result follows. Thus, for each $i \in \phi_{s_1}(\succ, P')$, $i \succ_{s_1} i_1$. Moreover, by *Pareto efficiency* of ϕ, all seats at s_1 are assigned under $\phi(\succ, P')$. Since $i_1 \in DA_{s_1}(\succ, P')$, this implies that there exists $i_2 \in \phi_{s_1}(\succ, P') \setminus DA_{s_1}(\succ, P')$. Thus, $i_2 \succ_{s_1} i_1$ and $s_1 = DA_{i_1}(\succ, P') \neq DA_{i_2}(\succ, P')$. By *fairness* of the DA, we conclude that $DA_{i_2}(\succ, P')\ P'_{i_2}\ \phi_{i_2}(\succ, P') = s_1$.

[13] If $DA_{i_1}(\succ, P') \neq \mu_{i_1} = s_1$, P_{i_1} is a profitable deviation for student i_1 in problem $(\succ, P')$.

We repeat this procedure for i_2, constructing P''_{i_2} based on P'_{i_2} and $DA_{i_2}(\succ, P')$. At each repetition, each student, assigned to her most preferred school in an earlier preference profile under the DA, and thus, under the $EADA$, is also assigned to her most preferred school. There are two cases: (1) Eventually, we reach a problem for which at least one student's priority is violated under ϕ but not under the $EADA$, which is a contradiction, and the result follows. (2) Due to a finite number of students and schools, the process implies an *improvement cycle*: there is a preference profile $\tilde{P}$ and a group of students $\{i_1, i_2, \ldots, i_m\}$ such that for each $k \in \{2, \ldots, m\}$,

$$DA_{i_k}(\succ, \tilde{P})\ \tilde{P}_{i_k}\ \phi_{i_k}(\succ, \tilde{P})\ = DA_{i_{k-1}}(\succ, \tilde{P}),$$

and

$$DA_{i_1}(\succ, \tilde{P})\ \tilde{P}_{i_1}\ \phi_{i_1}(\succ, \tilde{P})\ = DA_{i_m}(\succ, \tilde{P}).$$

This contradicts to ϕ being *Pareto efficient*, and the result follows.

Case 2: $\phi(\succ, P)$ Pareto dominates $DA(\succ, P)$. The matching $DA(\succ, P)$ is not *Pareto efficient* and thus, $E(\succ, P)$ *Pareto dominates* $DA(\succ, P)$. Since the students assigned to *underdemanded* schools under $DA(\succ, P)$ cannot be part of an *improvement cycle*, these schools have the same assignments under $\phi(\succ, P)$, $E(\succ, P)$ and $DA(\succ, P)$. Let $\mu \equiv DA(\succ, P)$ and let μ^k be the outcome of the TP in Step k. We show that, for each k, each student is assigned to a weakly better school under $\phi(\succ, P)$ than under μ^k. By *Pareto efficiency* of ϕ and the $EADA$, this implies that $\phi(\succ, P) = E(\succ, P)$.

Since $\phi(\succ, P)$ *Pareto dominates* μ, for each $i \in I$, $\phi_i(\succ, P)\ R_i\ \mu_i$. Suppose that in Step 1 of the TP, for some $i \in I$, $\mu^1_i\ P_i\ \phi_i(\succ, P)\ R_i\ \mu_i$. Since $\mu^1_i\ P_i\ \mu_i$, i is part of the *top priority cycle* in Step 1 of the TP. Let $i \to i_1 \to i_2 \to \ldots \to i_n \to i$ be this cycle. Let $\mu_i \equiv s$, and for each $m = 1, \ldots, n$, let $\mu_{i_m} \equiv s_m$. Since μ^1 is the outcome obtained by implementing this cycle, $\mu^1_i = \mu_{i_1} \equiv s_1$, $\mu^1_{i_n} = \mu_i \equiv s \equiv s_{n+1}$ and for each $m = 1, \ldots, n-1$, $\mu^1_{i_m} = \mu_{i_{m+1}} \equiv s_{m+1}$.

We construct a preference profile by consecutively constructing the preference order of each student in this *top priority cycle*. We start with student i.

Let

$$\tilde{S}_0 \equiv \mathcal{U}_s(P_i) \setminus \mathcal{U}_{s_1}(P_i) = \{s' \in S : s_1 \; R_i \; s' \; P_i \; s\}.$$

Let P_i' be the preference order constructed based on P_i and $\tilde{S}_0$, and $P^0 = (P_i', P_{-i})$ such that:

i. for each $s' \in \tilde{S}_0$ and $s'' \in S \setminus \tilde{S}_0$, $s' \; P_i' \; s''$,

ii. for each $s', s'' \in \tilde{S}_0$, $s' \; P_i \; s''$ implies $s' \; P_i' \; s''$,

iii. for each $s', s'' \in S \setminus \tilde{S}_0$, $s' \; P_i \; s''$ implies $s' \; P_i' \; s''$.

The preference order P_i' is an *upper-manipulation* of P_i at μ_i, and also an *upper-manipulation* of P_i at $\phi_i(\succ, P)$.

By (i) and (iii), $\mathcal{U}_s(P_i) = \mathcal{U}_s(P_i')$ and $\mathcal{L}_s(P_i) = \mathcal{L}_s(P_i')$. By Claim 4 (see the proof of Proposition 3), we conclude that $DA(\succ, P^0) = DA(\succ, P) = \mu$. Moreover, since $\mathcal{U}_s(P_i') = \mathcal{U}_s(P_i)$, the *top priority cycle* implemented in Step 1 of the TP is identical under P and P^0. Thus, i is assigned to s_1 in Step 1 of the TP under $(\succ, P^0)$. By (i) and (ii), s_1 is the most preferred school for i under P_i'. Thus, $E_i(\succ, P^0) = s_1$. Since ϕ is *upper-manipulation-proof*, $\phi_i(\succ, P) \; R_i \; \phi_i(\succ, P^0)$. Thus,

$$E_i(\succ, P^0) = s_1 \; P_i \; \phi_i(\succ, P) \; R_i \; \phi_i(\succ, P^0).$$

If $\phi(\succ, P^0)$ does not *Pareto dominate* $DA(\succ, P^0) = \mu$, then the result follows by Case 1 for problem $(\succ, P')$. Suppose $\phi(\succ, P^0)$ *Pareto dominates* $DA(\succ, P^0) = \mu$. Since $\mathcal{U}_s(P_i') = \mathcal{U}_s(P_i)$, the *underdemanded* schools in problems $(\succ, P)$ and $(\succ, P^0)$ under μ are the same. Since students assigned to *underdemanded* schools cannot be part of an *improvement cycle*, these students have the same assignment under $\phi(\succ, P^0)$ as they have under μ. By definition of the TP, i has the highest priority at s_1 among the students who are not assigned to an *underdemanded* school under μ and who prefer s_1 to their assignment under μ (†). Suppose there is a student j assigned to s_1 under $\phi(\succ, P^0)$ but not under μ. Then, j is not assigned to an *underdemanded* school under μ. Also, since $\phi(\succ, P^0)$ *Pareto dominates* $DA(\succ, P^0) = \mu$, j prefers s_1 to μ_j. By (†), this implies $i \succ_{s_1} j$. But then, i has *justified-envy* at $\phi(\succ, P^0)$ but not at $E(\succ, P^0)$, which contradicts to ϕ having *weakly less justified-envy* than the $EADA$. We conclude that no student who is not assigned to s_1 under μ is assigned to s_1 under $\phi(\succ, P^0)$. Thus, the assignment of s_1 is the same under $\phi(\succ, P^0)$ and μ.

We repeat this argument consecutively for each student who is part of the *top priority cycle* implemented in Step 1 of the TP under $(\succ, P)$.

For $m = 1, \ldots, n$, let

$$\tilde{S}_m \equiv \mathcal{U}_{s_m}(P_{i_m}) \setminus \mathcal{U}_{s_{m+1}}(P_{i_m}) = \{s' \in S : s_{m+1}\ R_{i_m}\ s'\ P_{i_m}\ s_m\},$$

and P'_{i_m} be the preference order, constructed based on P_{i_m} and $\tilde{S}_m$, such that:

i. for each $s' \in \tilde{S}_m$ and $s'' \in S \setminus \tilde{S}_m$, $s'\ P'_{i_m}\ s''$,

ii. for each $s', s'' \in \tilde{S}_m$, $s'\ P_{i_m}\ s''$ implies $s'\ P'_{i_m}\ s''$,

iii. for each $s', s'' \in S \setminus \tilde{S}_m$, $s'\ P_{i_m}\ s''$ implies $s'\ P'_{i_m}\ s''$,

and $P^m = (P'_i, P'_{i_1}, \ldots, P'_{i_m}, P_{-\{i, i_1, \ldots i_m\}})$.

We repeat the argument for P'_i and P^0 above also for P'_{i_m} and P^m consecutively for each $m = 1, \ldots, n$, and we obtain:

I. $DA(\succ, P^m) = \mu$.

II. The *underdemanded* schools in problem $(\succ, P^m)$ under $DA(\succ, P^m) = \mu$ are the same as in problem $(\succ, P^{m-1})$ under $DA(\succ, P^{m-1}) = \mu$, and by induction, the same as in problem $(\succ, P)$ under $DA(\succ, P) = \mu$.

III. The *top priority cycle* implemented in Step 1 of the TP under P^m is the same as the one under P^{m-1}, and by induction, the same as the one under P.

IV. The assignment of s_{m+1} is the same under $\phi(\succ, P^m)$ and μ.

By (IV), i is assigned to s under $\phi(\succ, P^n)$. By applying the same argument (†) and the fact that ϕ has *weakly less justified-envy* than the $EADA$, we conclude that the assignment of s_1 is the same under $\phi(\succ, P^n)$ and μ. Then, i_1 is assigned to s_1 under $\phi(\succ, P^n)$. By repeating the same argument consecutively, we conclude that, for each $m = 1, \ldots, n$, i_m is assigned to s_m under $\phi(\succ, P^n)$. But then, there exists an *improvement cycle* under $\phi(\succ, P^n)$, which is the *top priority cycle* implemented in Step 1 of the TP under $(\succ, P^n)$. This contradicts to *Pareto efficiency* of ϕ. We conclude that for each $j \in I$,

$$\phi_j(\succ, P)\ R_j\ \mu_j^1.$$

We next show that each student assigned to an *underdemanded* school under μ^1 is assigned to the same school under $E(\succ, P)$ and $\phi(\succ, P)$. We start with an *undemanded* school. Let i be a student such that there does not exist $j \in I$ such that $\mu_i^1\ P_j\ \mu_j^1$. Since for each $j \in I$, $E_j(\succ, P)\ R_j\ \mu_j^1$ and $\phi_j(\succ, P)\ R_j\ \mu_j^1$, $\mu_i^1 = E_i(\succ, P) = \phi_i(\succ, P)$. Moreover, if a student is assigned to a school which is demanded only by students assigned to *underdemanded* schools in μ^1 under $E(\succ, P)$, then so is she under $\phi(\succ, P)$.

We now consider μ^2. Suppose, towards a contradiction, that for some $i \in I$, $\mu_i^2\ P_i\ \phi_i(\succ, P)\ R_i\ \mu_i^1$. Let

$$\hat{S} \equiv \mathcal{U}_{\mu_i}(P_i) \setminus \mathcal{U}_{\mu_i^2}(P_i) = \{s \in S : \mu_i^2\ R_i\ s\ P_i\ \mu_i\}.$$

Let P_i'' be the preference order, constructed based on P_i and $\hat{S}$, and $P'' = (P_i'', P_{-i})$ such that:

i. for each $s' \in \hat{S}$ and $s'' \in S \setminus \hat{S}$, $s'\ P_i''\ s''$,

ii. for each $s', s'' \in \hat{S}$, $s' \; P_i \; s''$ implies $s' \; P''_i \; s''$,

iii. for each $s', s'' \in S \setminus \hat{S}$, $s' \; P_i \; s''$ implies $s' \; P''_i \; s''$.

The preference order P'_i is an *upper-manipulation* of P_i at μ_i, and and also an *upper-manipulation* of P_i at $\phi_i(\succ, P)$.

By Claim 4, $DA(\succ, P'') = DA(\succ, P) = \mu$. Thus, the cycles implemented in the first two steps of the TP in problem $(\succ, P)$ also exist when the TP is applied to problem $(\succ, P'')$. Thus, $E_i(\succ, P'') = \mu_i^2$. Since ϕ is *upper-manipulation-proof*, $\phi_i(\succ, P) \; R_i \; \phi_i(\succ, P'')$. Thus,

$$E_i(\succ, P'') = \mu_i^2 \; P_i \; \phi_i(\succ, P) \; R_i \; \phi_i(\succ, P'').$$

If $\phi(\succ, P'')$ does not *Pareto dominate* $DA(\succ, P'') = \mu$, then the result follows by Case 1 for problem $(\succ, P'')$. Suppose $\phi(\succ, P'')$ *Pareto dominates* $DA(\succ, P'')$. Since the TP selects the same cycles in the first step under both problems $(\succ, P)$ and $(\succ, P'')$, μ^1 is the matching obtained in the first step of the TP for problems $(\succ, P)$ and $(\succ, P'')$. As shown above, for each $j \in I$, $\phi_j(\succ, P'') \; R''_j \; \mu_j^1$. Then, each student who is assigned to an *underdemanded* school under μ^1 cannot be assigned to a better school under $\phi(\succ, P'')$. By definition of the TP, i has the highest priority at school μ_i^2 among the students who are not assigned to an *underdemanded* school under μ^1 and who prefer μ_i^2 to their assignment under μ^1. Thus, if there is a student assigned to μ_i^2 under $\phi(\succ, P'')$ but not under μ^1, then ϕ cannot have *weakly less justified-envy* than the $EADA$. We conclude that the assignment of μ_i^2 is the same under $\phi(\succ, P'')$ and μ^1. By repeating the same argument above, we have an *improvement cycle* under $\phi(\succ, P'')$. This contradicts to *Pareto efficiency* of ϕ. We conclude that for each $j \in I$,

$$\phi_j(\succ, P) \; R_j \; \mu_j^2.$$

By repeating the same argument for the following steps of the TP, we achieve the result. □

Appendix D Additional Results

Proposition 5. *The $EADA$ is both strongly t-justified-envy minimal and strongly p-justified-envy minimal in the class of Pareto efficient mechanisms.*

Proof. Suppose, towards a contradiction, that ϕ is *Pareto efficient* and has *weakly less p-justified-envy* than the $EADA$. Let $(\succ, P)$ be a problem where $E(\succ, P) \neq \phi(\succ, P)$. If $E(\succ, P) = DA(\succ, P)$, then $E(\succ, P)$ is stable. Since ϕ has *weakly less p-justified-envy* than the $EADA$ and $E(\succ, P)$ is stable, we have $\phi(\succ, P) = DA(\succ, P)$, implying that $E(\succ, P) = \phi(\succ, P)$, contradicting our starting supposition. Thus, $E(\succ, P) \neq DA(\succ, P)$. Since the $EADA$ is *p-justified-envy minimal* in the class of *Pareto efficient* mechanisms (Theorem 1 of Kwon and Shorrer (2020)), and ϕ has *weakly less p-justified-envy* than the $EADA$, we have $JEP^{E(\succ,P)}(\succ, P) = JEP^{\phi(\succ,P)}(\succ, P)$.

Let $(\succ, P')$ be the problem such that P' is obtained from P as follows: for each $(i, s) \in JEP^{E(\succ,P)}(\succ, P)$, i ranks s below s_0, while keeping the relative ranking of other schools the same. We have $E(\succ, P') = DA(\succ, P')$. Moreover, by the definition of the $EADA$, $E(\succ, P) = DA(\succ, P')$.

We now claim that $\phi(\succ, P) = DA(\succ, P')$. First, note that $DA(\succ, P') = E(\succ, P')$ is efficient in problem $(\succ, P')$. By definition of DA, in problem $(\succ, P')$, there is a unique efficient matching that has no justified envy, which is $DA(\succ, P')$. Therefore, in problem $(\succ, P)$, there is a unique efficient matching that has no blocking pairs apart from those in $JEP^{\phi(\succ,P)}(\succ, P)$, which is $DA(\succ, P')$. Thus, $\phi(\succ, P) = DA(\succ, P')$.

Recall that $E(\succ, P) = DA(\succ, P')$. Thus, $E(\succ, P) = \phi(\succ, P)$, a contradiction, showing the result for p-justified envy-freeness. By the same arguments, we can straightforwardly show the same result for strong t-justified-envy minimality as well. □

References

Abdulkadiroğlu, Atila, Yeon-Koo Che, Parag A. Pathak, Alvin E. Roth, and Olivier Tercieux (2020) "Efficiency, Justified Envy, and Incentives in Priority-Based Matching," *American Economic Review: Insights*, 2 (4), 425–42.

Abdulkadiroğlu, Atila and Aram Grigoryan (2021) "Efficient and Envy Minimal Matching," Working Paper.

Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth (2009) "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *American Economic Review*, 99, 1954–1978.

Abdulkadiroğlu, Atila and Tayfun Sönmez (2003) "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

Abraham, David J., Péter Biró, and David F. Manlove (2005) "'Almost Stable' Matchings in the Roommates Problem," in *International Workshop on Approximation and Online Algorithms*, 1–14, Springer.

Alcalde, José and Antonio Romero-Medina (2017) "Fair Student Placement," *Theory and Decision*, 83, 293–307.

Alva, Samson and Vikram Manjunath (2019) "Strategy-proof Pareto Improvement," *Journal of Economic Theory*, 181, 121–142.

Balinski, Michel and Tayfun Sönmez (1999) "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 84, 73–94.

Bogomolnaia, Anna and Hervé Moulin (2001) "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 100, 295–328.

Cerrone, Claudia, Claudia Cerrone, Yoan Hermstrüwer, and Onur Kesten (2022) "School Choice with Consent: An Experiment," *MPI Collective Goods Discussion Paper, No. 2022/2.*

Chen, Yan and Tayfun Sönmez (2006) "School Choice: An Experimental Study," *Journal of Economic Theory*, 127, 202–231.

Doğan, Battal and Lars Ehlers (2021a) "Minimally Unstable Pareto Improvements over Deferred Acceptance," *Theoretical Economics*, 16, 1249–1279.

——— (2021b) "Robust Minimal Instability of The Top Trading Cycles Mechanism," *American Economic Journal: Microeconomics*, forthcoming.

Dubins, Lester E. and David A. Freedman (1981) "Machiavelli and the Gale-Shapley Algorithm," *American Mathematical Monthly*, 88, 485–494.

Dur, Umut, A. Arda Gitmez, and Özgür Yılmaz (2019) "School Choice under Partial Fairness," *Theoretical Economics*, 14, 1309–1346.

Ehlers, Lars and Thayer Morrill (2020) "(Il)legal Assignments in School Choice," *Review of Economic Studies*, 87 (4), 1837–1875.

Gale, David and Lloyd S. Shapley (1962) "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9–15.

Hashimoto, Tadashi, Daisuke Hirata, Onur Kesten, Morimitsu Kurino, and M. Utku Ünver (2014) "Two Axiomatic Approaches to the Probabilistic Serial Mechanism," *Theoretical Economics*, 9 (1), 253–277.

Kesten, Onur (2010) "School Choice with Consent," *Quarterly Journal of Economics*, 125, 1297–1348.

Kesten, Onur and Morimitsu Kurino (2019) "Strategy-proof Improvements upon Deferred Acceptance: A Maximal Domain for Possibility," *Games and Economic Behavior*, 117, 120–143.

Kojima, Fuhito and Mihai Manea (2010) "Axioms for Deferred Acceptance," *Econometrica*, 78, 633–653.

Kwon, Hyukjun and Ran I. Shorrer (2020) "Justified-Envy Minimal Mechanisms in School Choice," Working Paper.

Mennle, Timo and Sven Seuken (2021) "Partial Strategyproofness: Relaxing Strategyproofness for the Random Assignment Problem," *Journal of Economic Theory*, 191, 147–151.

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

Tang, Qianfeng and Jingsheng Yu (2014) "A New Perspective on Kesten's School Choice with Consent Idea," *Journal of Economic Theory*, 154 (0), 543 – 561.

Tang, Qianfeng and Yongchao Zhang (2021) "Weak Stability and Pareto Efficiency in School Choice," *Economic Theory*, 71, 533–552.

Troyan, Peter, David Delacrétaz, and Andrew Kloosterman (2020) "Essentially Stable Matchings," *Games and Economic Behavior*, 120, 370–390.