Regret-Free Truth-Telling in School Choice with Consent*

Yiqiu Chen Markus Möller†

July 26, 2022

Abstract

The *Efficiency Adjusted Deferred Acceptance Rule (EDA)* is a promising candidate mechanism for public school assignment. A potential drawback of EDA is that it could encourage students to game the system since it is not strategy-proof. However, to successfully strategize, students typically need information that is unlikely to be available to them in practice. We model school choice under incomplete information and show that EDA is regret-free truth-telling, which is a weaker incentive property than strategy-proofness and was introduced by Fernandez (2020). We also show that there is no efficient matching rule that weakly Pareto dominates a stable matching rule and is regret-free truth-telling. Note that the original version of EDA by Kesten (2010) weakly Pareto dominates a stable matching rule, but it is not efficient.

Keywords: School Choice, Matching, Efficiency Adjusted Deferred Acceptance, Regret, Manipulation, Stable-Dominating.

JEL Codes: C78, D81, D82, I20.

---

*We thank especially our advisor, Alexander Westkamp. We are grateful to Christoph Schottmüller, Marcelo Ariel Fernandez, Kevin Breuer and Marius Gramb for helpful comments. All errors remain our own. Chen and Möller acknowledge financial support from the Cologne Graduate School of Management, Economics and Social Sciences (University of Cologne). Chen and Möller are funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germanys Excellence Strategy – EXC 2126/1– 390838866.

†University of Cologne - E-mail: chen@wiso.uni-koeln.de (Chen) moeller@wiso.uni-koeln.de (Möller)

1 Introduction

Efficiency and fairness are incompatible in the school choice problem.[1] The *Efficiency Adjusted Deferred Acceptance Rule (EDA)* (Kesten, 2010) elegantly circumvents this incompatibility by allowing students to give their consent to relax the fairness constraint. Its desirable features made EDA a candidate for school assignment in Belgium's Flanders region in 2019 (Cerrone et al., 2022). However, EDA belongs to the class of *stable dominating (matching) rules* (Alva and Manjunath, 2019a) of which no candidate is strategy-proof except the *Student-Proposing Deferred Acceptance Rule (DA)* (Alva and Manjunath, 2019b).[2,3] To address possible incentive issues with EDA, we examine whether it satisfies an incentive criterion by Fernandez (2020) which is weaker than strategy-proofness and is based on participants' wish to avoid regret.

We employ a many-to-one school choice model with consent (Kesten, 2010) under incomplete information, where students can reconsider their admission chances for alternative reports through an observational structure based on the cutoff terminology. We express each school's individual priorities over students in the form of scores and for each school, the cutoff is the lowest score among all students that have been admitted to that school. Once the final matching has been determined, each student makes an observation that consists of the final matching and each school's cutoff and can then draw inferences about the set of market unknowns that are consistent with

[1]A student has justified envy at a matching, if there exists a lower prioritized student assigned to a school and the corresponding school is preferred to her assignment (Abdulkadiroğlu and Sönmez, 2003). A matching is *fair* if no justified envy exists and a matching rule is fair if it only produces matchings which are fair. The trade-off between efficiency and fairness follows from Balinski and Sönmez (1999).

[2]A matching rule is *stable* if it produces outcomes which are *fair*, *individually rational* and *non-wasteful*. A matching is non-wasteful if there is no object that is unassigned although there is an agent that prefers it over her assignment. A matching is individually rational if no agent prefers her outside option over her final assignment. A stable dominating rule always implements a matching that weakly Pareto dominates a stable matching (Alva and Manjunath, 2019a).

[3]Strategy-proofness requires that it is a weakly dominant strategy for students to report their true preferences. DA was introduced by Gale and Shapley (1962) and shown to be strategy-proof by Dubins and Freedman (1981) and Roth (1982). For related results regarding the incompatibility of strategy-proofness with rules that Pareto dominate DA, see also Abdulkadiroğlu et al. (2009), Erdil and Ergin (2008) or Kesten (2010).

her observation. Our choice of a student's unknowns is motivated by characteristics common in the context of public school assignment, and includes other students' scores and their reported preferences. Specifically, it is common in practical applications that students' scores are based on proximity, walk-zone areas, sibling-status and other socioeconomic variables. The composition of scores is usually public information, whereas accurate information on other students' scores and reported preferences will generally be covered by privacy protection.

In this framework, we adopt an incentive notion by Fernandez (2020) that is based on regret. A student *regrets* her report at an observation if she finds another report, which does not assign her worse for all market unknowns compatible with the observation and assigns her strictly better for some of the compatible market unknowns. A rule is *regret-free truth-telling* if no student ever regrets reporting her preferences truthfully.

The main finding of this paper is that EDA is *regret-free truth-telling* (Theorem 1) and that under EDA, truth-telling is the unique option which never leads to regret (Proposition 2). We thus provide an appropriate statement for the intuition that truth-telling can be a focal strategy under EDA and contribute to the strand of literature that outlines the many desirable features of EDA for practical implementation. Note that we assume that students make their inferences subject to the uncertainty and unobservability of other students' consents. The just described uncertainty plays a key role in the proof of Theorem 1.

We also study stable dominating rules without consent decisions. Under these rules, students indicate only their preference rankings over schools, so there is no uncertainty about the consent decisions of other students. As argued by Alva and Manjunath (2019a) stable dominating rules without consent decisions address the efficiency and fairness trade-off as follows: Students who complain that they experienced justified envy under a matching can always be offered a corrective, namely the stable matching which the implemented matching Pareto dominates. Since the corrective makes all students, including the students who complain, weakly worse off, it does not pay off to complain.

We show that there is a stable dominating rule without consent decisions which is not equivalent to DA and which is regret-free truth-telling (Proposition 3). However, we also show that stable dominating rules which are regret-free truth-telling cannot be efficient (Theorem 2). Stable dominating rules which are efficient contain some interesting candidates for practical applications such as a version of EDA that improves to the efficiency frontier without students' consents. Note that the original version of EDA for which Theorem 1 is satisfied, is not efficient since it respects improvements on efficiency only with students' consents.

Related Literature

To our knowledge, Fernandez (2020) is the first to introduce regret-based incentives in the matching literature.[4] In marriage markets, Fernandez (2020) shows that truth-telling is the unique regret-free strategy under DA for both men and women and that DA is the unique regret-free truth-telling rule among a subclass of stable rules which are called quantile stable rules.[5] Fernandez (2020) sheds light on college admissions problems. He shows that the student-proposing variant of DA is regret-free truth-telling. However, under the college-proposing variant of DA, being truthful does not need to be free of regret for colleges. The key differences of our work to that of Fernandez (2020) is that in our contribution only the students are strategic. Moreover, whereas in Fernandez (2020) participants only observe the realized matching, students in our model additionally observe cutoffs.

[4] Regret-based incentives have a long tradition in economic theory. For instance, in auction theory, regret-based incentives of bidders in first-price auctions have been studied by Filiz-Ozbay and Ozbay (2007) and Engelbrecht-Wiggans (1989). For a more detailed discussion we refer to Fernandez (2020). See Gilovich and Medvec (1995) and Zeelenberg and Pieters (2007) for psychological treatments of regret.

[5] For more information on quantile stable rules we refer to Teo and Sethuraman (1998), Klaus and Klijn (2006), or Chen et al. (2014).

This paper mainly contributes to the literature that deepens the understanding of EDA's incentive properties. Our results complement those of Troyan and Morrill (2020), who show that for cognitively limited participants beneficial misreporting under EDA is not *obvious* in the following sense: a profitable misreport is an *obvious manipulation* if the best-case outcome of the misreport is better than the best-case outcome of telling the truth or, if the worst-case outcome of the misreport is better than the worst-case outcome of telling the truth. The main difference between our work and that of Troyan and Morrill (2020) concerns the source of uncertainty that students face. A profitable misreport is obvious if it is easy to recognize for students whose knowledge on the matching rule is imperfect, given that these students have full access to the scores of other students. That is, non-obvious manipulability is mainly driven by participants' limited understanding of the matching rule. By contrast, students in our model know how the matching rule works and our results are driven by students' incomplete access to the scores of other students. Notably, the positive result of Troyan and Morrill (2020) covers both EDA and stable dominating rules, where we reach a negative result for efficient stable dominating rules.

Previous results on EDA's incentive properties are inspired by the theoretical benchmark for low information environments from Roth and Rothblum (1999) and Ehlers (2008). Kesten (2010) studies Bayesian incentives of EDA in a setting where it is common knowledge that students' preferences over schools are ordered into shared quality classes and students' beliefs on how other students order schools within each quality class are symmetrically distributed. Kesten (2010) shows that if other students submit their true preferences, then truth-telling stochastically dominates any other strategy. The key difference to our model is that we do not specify any prior probability distribution regarding the beliefs or distribution on other participants' preferences and thus do not impose any symmetry assumptions or correlation of preferences over schools. Thus, in contrast to the approach of Kesten (2010) our information environment follows the "Wilson doctrine" (Wilson, 1987).

Related to our work are also some more recent findings on EDA's incentive features. Reny (2021) shows that under EDA, truth-telling is a maxmin optimal strategy for students that do not know other students' preferences. Decerf and Van der Linden (2021) find that rules that Pareto dominate DA are harder to manipulate than the well-known Boston mechanism. Finally, a recent experiment on manipulation under EDA by Cerrone et al. (2022) revealed that different variants of EDA yield higher rates of truth-telling than DA in environments with strategic uncertainty, complete information about the primitives and given that students are not allowed to truncate.

More generally, the theoretical literature on EDA is growing rapidly as well. Tang and Yu (2014), Ehlers and Morrill (2020), Bando (2014) and Dur et al. (2019) have recently developed tractable alternatives to Kesten's initial formulation of EDA. Ehlers and Morrill (2020) generalize EDA to a school choice model where school priorities take the form of flexible choice functions and Kwon and Shorrer (2019) propose a version of EDA for organ exchange. EDA also manages to satisfy some reasonable weaker alternatives to fairness in the sense of Abdulkadiroğlu and Sönmez (2003), including for example, guaranteed selection of *essentially stable matchings* (Troyan et al., 2020), *priority-neutral* matchings (Reny, 2021) and *legal* matchings (Ehlers and Morrill, 2020).

Further, our paper relates to the line of literature that uses the cutoff terminology in school choice models. Most prominent in this regard is Azevedo and Leshno (2016) who characterize stable matchings in terms of cutoffs in a continuum school choice model. They show that cutoffs take the form of market-clearing prices that equalize supply and demand and can be used to perform comparative statics with respect to schools' incentives to invest in quality. When used to characterize stable matchings, cutoffs usually take the form of a guarantee for participants to be admitted at schools. In our framework, final assignments may not correspond to stable matchings. Therefore, the cutoffs do not necessarily provide a student with information about whether she will be admitted at a desired school. Moreover, in our model the cutoffs are incorporated into students' strategic reasoning.

Finally, this work also adds to the literature examining the impact of behavioral biases on decision making in school choice. Meisner and von Wangenheim (2021) and Dreyfuss et al. (2019) show that students not playing truthfully under the student-proposing variant of DA can be explained by students being loss averse. The influence of students' overconfidence is examined by Pan (2019).

The rest of this paper is organized as follows. We introduce the basic model and EDA in Section 2. We model the informational environment and adopt regret-free truth-telling in Section 3. In Section 4, we present our main results. Our analysis regarding stable dominating rules without consent option is provided in Section 5. In Section 6, we give a brief discussion of how our key assumptions influence the results. Finally, Section 7 gives a short conclusion. The Appendix contains most of our proofs.

2 The Model

There is a finite set of students I and a finite set of schools S. Each school $s \in S$ has a fixed capacity q_s and we collect the capacities in $q = (q_s)_{s \in S}$. We add a common outside option $s_\emptyset$ for students which has infinite capacity.

Each school $s \in S$ has a vector of scores $g^s = \{g_i^s\}_{i \in I}$, where $g_i^s \in (0, 1)$ is i's score at s. We assume that $g_i^s \neq g_j^s$ for any $i, j \in I$ and any $s \in S$, and we say that for each pair of students $i, j \in I$, i has higher priority at s than j if and only if $g_i^s > g_j^s$. That is, for each school s, the school's scores induce a strict priority ranking over I.[6] For each $i \in I$, let $g_i = \{g_i^s\}_{s \in S}$ be the vector of scores assigned to student i. Let a score structure $g = (g_i)_{i \in I}$ be a collection of scores for each student and let $g_{-i} = (g_j)_{j \in I \setminus \{i\}}$ be a collection of scores for students in $I \setminus \{i\}$. Moreover, set $\mathcal{G}_I$ as the domain of all possible score structures and $\mathcal{G}_{-i}$ as the domain of all score structures for students

[6] The incomplete information framework we introduce in Section 3 allows students to draw inferences about their admission chances. Our formulation of scores will then ensure that a student typically cannot infer her exact rank on a school's priority list just on the basis of her own score.

other than i.

For each student $i \in I$, let $\succ_i$ be a strict preference relation over $S \cup \{s_\emptyset\}$. The corresponding weak preference relation of $\succ_i$ is denoted by $\succeq_i$.[7] Let $\mathcal{P}$ denote the set of all possible strict preference relations over $S \cup \{s_\emptyset\}$. For any $\succ_i \in \mathcal{P}$, a school s is acceptable to i if $s \succ_i s_\emptyset$ and unacceptable if it is not acceptable. A preference profile $\succ = (\succ_i)_{i \in I}$ is a realization of $\mathcal{P}$ for each $i \in I$ and $\succ_{-i} = (\succ_j)_{j \in I \setminus \{i\}}$ is a preference profile for students in $I \setminus \{i\}$. We define $\mathcal{P}_I$ as the domain of all preference profiles and $\mathcal{P}_{-i}$ as the domain of all preference profiles for students in $I \setminus \{i\}$.

A *matching* $\mu : I \to S \cup \{s_\emptyset\}$ is a function such that for each $s \in S$, $|\mu^{-1}(s)| \leq q_s$. Given any μ, we set $\mu_i = \mu(i)$ as the assignment of i and $\mu_s = \mu^{-1}(s)$ as the set of students assigned to s. Denote the set of all possible matchings by $\mathcal{M}$.

In the following, fix any $\succ \in \mathcal{P}_I$. We say a matching μ *weakly Pareto dominates* another matching μ' if for all $i \in I$, $\mu_i \succeq_i \mu'_i$. A matching μ *Pareto dominates* μ' if μ weakly Pareto dominates μ' and for some $j \in I$, $\mu_j \succ_j \mu'_j$. A matching μ is *Pareto efficient* if there does not exist another matching μ' which Pareto dominates μ.

We now introduce two fairness notions, where we start with the well-known notion by Abdulkadiroğlu and Sönmez (2003). Given a matching μ, student i has *justified envy* towards student j at school μ_j under μ if $\mu_j \succ_i \mu_i$ and $g_i^{\mu_j} > g_j^{\mu_j}$. A matching μ is *fair* if no student has justified envy at μ. A matching μ is *individually rational* if for each student the assigned school is acceptable to her. A matching μ is *non-wasteful* if there does not exist a student i and a school s, such that $s \succ_i \mu_i$ and $|\mu_s| < q_s$. A matching μ is *stable* if it is fair, individually rational and non-wasteful.

We also consider a weaker fairness notion that was introduced by Kesten (2010). The notion takes students' willingness to consent for being exposed to justified envy into account. For each student i, the consent is parameterized by a binary variable $c_i \in \{0, 1\}$, where $c_i = 1$ means that i consents and $c_i = 0$ means that i does not consent. We say a matching μ *violates the priority* of student i given c_i if $c_i = 0$

[7] That is, for all $s, s' \in S$, $s \succeq_i s'$ if either $s \succ_i s'$ or $s = s'$.

and if there exists another student $j \in I$ such that i has justified envy towards j at μ. Let $c = (c_i)_{i \in I}$ be a consent profile and let $\mathcal{C}_I$ be the domain of all consent profiles. Denote a consent profile of students other than i by $c_{-i} = (c_j)_{j \in I \setminus \{i\}}$ and the respective domain by $\mathcal{C}_{-i}$. Given a matching μ, a profile of preferences $\succ$ and a consent profile c, we say that a matching is *fair with consent* if there exists no student whose priority is violated at μ.

We call a collection $(I, S, q, g, \succ, c)$ a *school choice problem with consent* (or simply a *problem*). Throughout the main body, we fix a problem $(I, S, q, g, \succ, c)$. A *report* of student i is pair $(\succ'_i, c'_i) \in \mathcal{P} \times \{0, 1\}$ and a report profile is a pair $(\succ', c') \in \mathcal{P}_I \times \mathcal{C}_I$. Analogously, let $(\succ'_{-i}, c'_{-i}) \in \mathcal{P}_{-i} \times \mathcal{C}_{-i}$ be a report profile of students except i.

A *(matching) rule* $f : \mathcal{G}_I \times \mathcal{P}_I \times \mathcal{C}_I \to \mathcal{M}$ maps any triple of a score structure, preference profile and consent profile into a matching. Given a report profile $(\succ, c)$ and a score structure g, let the outcome of f be $f(g, \succ, c)$ and, for each $i \in I$, let $f_i(g, \succ, c)$ denote student i's respective assignment. If the rule does not take consent decisions into consideration, we write $f(g, \succ)$ instead of $f(g, \succ, c)$. A rule f is *Pareto efficient* if each outcome of the rule is Pareto efficient. Similarly, a rule is *stable* if it produces a stable matching for any problem. A rule f is *stable dominating* (Alva and Manjunath, 2019a) if for any problem $(I', S', q', g', \succ', c')$ the matching $f(g', \succ', c')$ weakly Pareto dominates a matching $\mu \in \mathcal{M}$ given $\succ'$, where μ is stable with respect to $(g', \succ')$.

We proceed with the description of two incentive notions for students. A matching rule f is *consent-invariant* if $f_i(g, \succ, (c_i, c_{-i})) = f_i(g, \succ, (c'_i, c_{-i}))$ for all i and all c_i, c'_i. That is, each student's assignment is independent of her own consent decision. Note that the rules studied in this paper are all consent-invariant. A matching rule f is *strategy-proof* if $f_i(g, (\succ_i, \succ_{-i}), c) \succeq_i f_i(g, (\tilde{\succ}_i, \succ_{-i}), c)$ for all i and all $\tilde{\succ}_i \in \mathcal{P}$. That means, for each student, reporting her true preferences is weakly better than reporting untruthfully regardless of other students' reports.

2.1 EDA

In this subsection, we present Kesten's EDA along with our first result. EDA inputs a report profile $(\succ, c)$ and produces an outcome where no student who decided not to consent experiences justified envy. EDA is stable dominating and essentially iteratively runs DA presented in Appendix A. Specifically, in the DA application process, a pair $(i, s) \in I \times S$ is an *interrupting pair* at step t' if (1) student i is tentatively accepted by s at some step t and is then rejected by s at the later step t' and; (2) another student is rejected by s at some step t^* with $t \leq t^* < t'$. Hereinafter, we refer to i as an *interrupter* for s at step t'. The formal description of the algorithm which induces EDA as in Kesten (2010) is provided below, while the alternative *Top Priority Algorithm* (Dur et al., 2019) used in most of the proofs can be found in Appendix A. Given any input report profile, EDA yields the outcome via the following procedure:

Round 0 Run DA.

Round k, $k \geq 1$ Consider the application process of DA in Round $k-1$. If there are interrupting pairs in which the interrupter consents, find the last step of this process where a consenting interrupter is rejected by the school for which she is an interrupter. At that step, collect all interrupting pairs with a consenting interrupter. For each collected pair (i, s), remove s from i's input preferences of round $k-1$ and keep the relative ranking of all other schools as before. For all other students, keep their input preferences the same as in round $k-1$. Then, run DA with the updated preference profile and proceed to Round $k+1$. If there are no interrupting pairs with a consenting interrupter, the algorithm terminates with the DA outcome of Round $k-1$.

We now move to our discussion on EDA's incentive properties which is known to be consent-invariant but not strategy-proof (Kesten, 2010). Our first result, Proposition 1, states that a certain class of deviations of a student does not affect her own assignment.

For any preference relation $\succ_i \in \mathcal{P}$ and school $s \in S$, let the weak lower contour set of $\succ_i$ with respect to s be $L_s^{\succ_i} = \{s' \in S \mid s \succeq_i s'\}$.

Proposition 1. *If $EDA(g, \succ, c) = \mu$ and $\tilde{\succ}_i \in \mathcal{P}$ is such that for all $s, s' \in L_{\mu_i}^{\succ_i}$, $s \succ_i s'$ only if $s \tilde{\succ}_i s'$, then $EDA_i(g, (\tilde{\succ}_i, \succ_{-i}), c) = \mu_i$.*

Proof. See Appendix B. □

In words, Proposition 1 shows that if a student's deviation from her baseline report keeps the order of the schools in the lower contour set with respect to the baseline assignment, then it yields the same outcome for the deviating student. The set of deviations considered in Proposition 1 is a subset of monotonic transformations at the student's baseline assignment. Formally, $\succ_i'$ is a *monotonic transformation* of $\succ_i$ at $s \in S \cup \{s_\emptyset\}$ if $s' \succ_i' s$ implies that $s' \succ_i s$. As will be evident from Section 4, Proposition 1 cannot be generalized to hold for all monotonic transformations at μ_i.

3 Regret in School Choice

In this section, we introduce the informational environment and regret-based incentives. We first describe the students' information and impose an observational structure. Assume that before submitting the report, each student i knows (I, S, q, g_i) and the matching rule f. After assignments have been determined by f, each student observes the matching and the cutoff at each school, i.e. the lowest score among all applicants matched to the school. More formally, given a report profile $(\hat{\succ}, \hat{c})$, student i observes $\mu = f(g, \hat{\succ}, \hat{c})$ and for each school $s \in S \cup \{s_\emptyset\}$, she observes $\pi_s(\mu, g) = \min_{i \in \mu_s} g_i^s$ when $|\mu_s| = q_s$ and $\pi_s(\mu, g) = 0$ otherwise. Let $\pi(\mu, g) = \{\pi_s(\mu, g)\}_{s \in S \cup \{s_\emptyset\}}$ and let an *observation* of student i be captured by $(\mu, \pi(\mu, g))$.

Next, define any triple $(g'_{-i}, \succ'_{-i}, c'_{-i}) \in \mathcal{G}_{-i} \times \mathcal{P}_{-i} \times \mathcal{C}_{-i}$ as a *scenario* for student i. If i submits $(\hat{\succ}_i, \hat{c}_i)$ and observes $(\mu, \pi(\mu, g))$, then scenario $(g'_{-i}, \succ'_{-i}, c'_{-i})$ is *plausible* if $\pi(\mu, g) = \pi(\mu, (g_i, g'_{-i}))$ and $f((g_i, g'_{-i}), (\hat{\succ}_i, \succ'_{-i}), (\hat{c}_i, c'_{-i})) = \mu$. The set of all plausible scenarios for student i is her *inference set* $\mathcal{I}(\mu, \hat{\succ}_i, \hat{c}_i)$. Moreover, for student $i \in I$ who reports $(\hat{\succ}_i, \hat{c}_i)$ to f, let

$$\mathcal{M}|_{(\hat{\succ}_i,\hat{c}_i)} = \{\mu \in \mathcal{M} \mid \exists(\succ'_{-i}, c'_{-i}) \in \mathcal{P}_{-i} \times \mathcal{C}_{-i} : f(g, (\hat{\succ}_i, \succ'_{-i}), (\hat{c}_i, c'_{-i})) = \mu\}$$

be the set of matchings that could be *observed* by student i. Note that g is fixed in $\mathcal{M}|_{(\hat{\succ}_i,\hat{c}_i)}$, since it is a primitive of the market and independent of the report profile.

Having defined our observational structure, we are ready to introduce the notions of regret and regret-free truth-telling adopted from Fernandez (2020). Recall that all matching rules we study are consent-invariant. To simplify our notation, we therefore define regret with a fixed consent decision for the student under consideration. Note, however, that for rules which are not consent-invariant one may define regret with respect to a pair of a consent decision and a preference ranking.

Definition 1. Fix consent decision $\hat{c}_i$. Student i *regrets* submitting $\hat{\succ}_i$ at $\mu \in \mathcal{M}|_{(\hat{\succ}_i,\hat{c}_i)}$ through $\succ^*_i$ under f if

1. $\forall(g'_{-i}, \succ'_{-i}, c'_{-i}) \in \mathcal{I}(\mu, \hat{\succ}_i, \hat{c}_i)$: $f_i((g_i, g'_{-i}), (\succ^*_i, \succ'_{-i}), (\hat{c}_i, c'_{-i})) \succeq_i \mu_i$
2. $\exists(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \hat{\succ}_i, \hat{c}_i)$: $f_i((g_i, \tilde{g}_{-i}), (\succ^*_i, \tilde{\succ}_{-i}), (\hat{c}_i, \tilde{c}_{-i})) \succ_i \mu_i$.

In words, a student regrets her report at an observation if there is an alternative report which guarantees her a weakly better assignment in all plausible scenarios and realizes a strict improvement in at least one plausible scenario.

Definition 2. Fix consent decision $\hat{c}_i$. A report $\hat{\succ}_i$ is *regret-free* under f if there does not exist a pair $(\mu, \succ^*_i) \in \mathcal{M}|_{(\hat{\succ}_i,\hat{c}_i)} \times \mathcal{P}$ such that i regrets $\hat{\succ}_i$ at μ through $\succ^*_i$.

That is, a regret-free report ensures that regardless of the realized observation, the student does not regret her report.

We only consider matching rules that are invariant in the unacceptable set and define reported preferences as *truth-telling* if they differ from a student's true preferences only in the order within the unacceptable set.

Definition 3. A matching rule f is *regret-free truth-telling* if for each problem and for each student, truth-telling is regret-free under f.

Strategy-proofness is stronger than regret-free truth-telling. That is, once truth-telling is weakly dominant under a rule, it is regret-free. However, the converse is not true. Specifically, strategy-proofness means that truth-telling is the weakly best option under *any* scenario, whereas regret-freeness only needs that, given a students' observation, no other report weakly dominates the truth in all *plausible scenarios*.

4 Main Results

In this section, we present our main result. We show that a student can avoid regret under EDA if she submits her true preferences (Theorem 1) and that there is no other reporting behavior that provides the same guarantee (Proposition 2).

Theorem 1. *EDA is regret-free truth-telling.*

Proof. See Appendix C. □

The following exposition provides an overview of the main arguments used in the formal proof. Fix any student $i \in I$, suppose that she reports her true preferences $\succ_i$ and she observes $(\mu, \pi(\mu, g))$. Then, any misreport $\tilde{\succ}_i$ can be interpreted as a combination of the following types of variations, where relative to $\succ_i$:

$(A1)$ for all $s, s' \in S$, $s \succ_i s'$ and $s' \tilde{\succ}_i s$ only if $s \in S \setminus L^{\succ_i}_{\mu_i}$;

$(A2)$ there exists $s' \in S$ such that $\mu_i \succ_i s'$ and $s' \tilde{\succ}_i \mu_i$, or;

$(A3)$ there exists $s, s' \in L^{\succ_i}_{\mu_i}$ such that $s, s' \in L^{\tilde{\succ}_i}_{\mu_i}$, $s \succ_i s'$ and $s' \tilde{\succ}_i s$.

Type $(A1)$ involves all variations relative to $\succ_i$ which keep the same ranking of all schools that are truly less preferred to μ_i. Type $(A2)$ considers the misreports which rank some schools that are truly less preferred to μ_i as more preferred and $(A3)$ considers the misreports which alter the rankings among the schools that are truly less preferred to μ_i.

First note that any variation $\tilde{\succ}_i$ of type $(A1)$ relates to Proposition 1. If $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i})$ is plausible, then we have $EDA((g_i, \tilde{g}_{-i}), (\tilde{\succ}_i, \tilde{\succ}_{-i}), (c_i, \tilde{c}_{-i})) = \mu$ and we can apply Proposition 1 to obtain $EDA_i((g_i, \tilde{g}_{-i}), (\tilde{\succ}_i, \tilde{\succ}_{-i}), (c_i, \tilde{c}_{-i})) = \mu_i$.

Next, let student i choose a misreport $\tilde{\succ}_i$ that contains variations of type $(A2)$ and let $\tilde{S} = \{s' \in S \mid \mu_i \succ_i s' \text{ and } s' \tilde{\succ}_i \mu_i\}$. The key arguments in the proof can roughly be divided into two categories: The submission of $\tilde{\succ}_i$ either would not have effectively influenced the assignment process at all, meaning i's assignment remains μ_i; or there is at least one plausible scenario in which the student is finally assigned to some $s^* \in \tilde{S}$. Here, we discuss the latter and more interesting case. The starting point of our argument is to construct a plausible scenario $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i})$ where i is assigned to s^* under $DA((g_i, \tilde{g}_{-i}), (\tilde{\succ}_i, \tilde{\succ}_{-i}))$. Then, we show that either the potential improvements that involve i cannot be realized because the consent of a student is missing; or there is no student who prefers s^* to her assignment under $DA((g_i, \tilde{g}_{-i}), (\tilde{\succ}_i, \tilde{\succ}_{-i}))$. The key challenge is to construct plausible scenarios for which one can isolate such a school $s^* \in L^{\mu_i}_{\succ_i}$ since the selection varies with the properties of the observed cutoffs.

Intuitively, given strategy-proofness of DA, what would allow student i to benefit from misreporting under EDA is that relative to the process under truth, (1) i's application at s^* creates a lastly rejected interrupting pair and (2) the created interrupter consents. In this case, the induced inefficiency under DA may lead student i to improve upon s^* to some school preferred to μ_i under EDA. However, there always exists a plausible scenario for student i where (1) or (2) cannot be satisfied, under which i is assigned to s^* and worse off compared to μ_i. Section 6 discusses in more detail in which cases the uncertainty about other students' consent decisions is needed for the result.

Also note that the intuition and the identified key challenges are not present in the analysis of Fernandez (2020) who shows that DA is regret-free truth-telling in the marriage problem. In Fernandez (2020) agents do not have the option to consent, do not observe cutoffs and the property is proven for the side that receives the applications under DA. Essentially, Fernandez (2020) shows that a preference profile for which the set of stable matchings is a singleton is always plausible. This means that i's observed assignment is already her best achievable stable assignment and student i may be worse off by misreporting. Since EDA is not stable and, in particular, the cutoffs may reveal the instability of the matching to the observing student, the argument by Fernandez (2020) is not applicable here.

Finally, suppose that the misreport $\tilde{\succ}_i$ contains variations of type ($A3$). The key argument for such a misreport is similar to that for ($A2$): By submitting $\tilde{\succ}_i$, student i faces the possibility to be assigned to a less preferred school s^* whose order is permuted in $\tilde{\succ}_i$ and for which there is no student who prefers s^* to her assignment under $DA((g_i, \tilde{g}_{-i}), (\tilde{\succ}_i, \tilde{\succ}_{-i}))$. However, different from ($A2$), here the target school s^* still ranks below μ_i on $\tilde{\succ}_i$ and its identification depends on its relative position to i's DA assignment on $\succ_i$ and $\tilde{\succ}_i$. This difference brings an additional challenge to the proof that is also absent in Fernandez (2020). While for ($A2$) it is enough to consider a plausible scenario where under truth-telling, i was already assigned to μ_i under DA, for ($A3$) we need to construct a scenario where under truth-telling, the updating procedure of EDA improves student i from some school to μ_i.

Our final result in this section shows that truth-telling is the unique regret-free choice under EDA.

Proposition 2. *For any non-truthful report, there exists an observation at which the student regrets it through truth-telling.*

Proof. See Appendix D. □

At first glance, it might appear that Proposition 1 and Proposition 2 are in conflict with each other. However, Proposition 1 only implies that a certain class of misreports does not change the student's assignment when we fixed an observation that follows from her true preferences. In Proposition 2, however, the observation is not fixed. Instead, we show that given any non-truthful report, we can find a corresponding observation, such that truth-telling guarantees weakly better assignments in all plausible scenarios.

As an intuition for Proposition 2 note that for every misreport there must exist a pair, say school s and $\tilde{s}$, which compared to the truth, reverse their rankings. Let student i prefer s to $\tilde{s}$ under truth. Now suppose that upon submission of the misreport, she is assigned to $\tilde{s}$ while a seat at s is vacant. Note that the vacant seat at s allows i to infer that the truth would have guaranteed her at worst s. As a result, she will regret not having been truthful. The key step in the proof is to construct an observation of the type just described for any misreport.

5 Stable Dominating Rules without Consent Decisions

In this section, we focus on stable dominating rules without consent decisions. That is, unlike under the version of EDA examined in the previous section, consent decisions are not reported for this class of stable dominating rules. This also means that students face no uncertainty regarding the consent decisions of other students. Accordingly, we modify the elements of the basic framework presented in Section 3 to reflect the removal of the consent decisions. To exemplify this point, let a scenario for student i reduce to a pair $(\tilde{g}_{-i}, \tilde{\succ}_{-i}) \in \mathcal{G}_{-i} \times \mathcal{P}_{-i}$ and denote i's inference set with $\mathcal{I}(\mu, \succ_i)$.

For our main negative result, we refine the set of stable dominating rules to consider only candidates which are efficient. A matching rule is *efficient stable dominating* if it is stable dominating and Pareto efficient. Since efficient stable dominating rules are stable dominating, it follows from Alva and Manjunath (2019b) that none of the them is strategy-proof. As we will show below, for efficient stable dominating rules, also regret-free truth-telling cannot be satisfied.

Theorem 2. *No efficient stable dominating rule is regret-free truth-telling.*

The proof below is constructive. We provide a problem with $|S| = 2$ and $|I| = 3$, and show that a student regrets submitting her true preferences under any efficient stable dominating rule. We only need small adjustments in the construction to apply the basic argument to any market with $|S| \geq 2$ and $|I| \geq 3$.

Proof. Consider a problem $(I, S, q, g, \succ)$ with two schools $S = \{s_1, s_2\}$ with capacities $q_{s_1} = q_{s_2} = 1$ and three students $I = \{i_1, i_2, i_3\}$. Suppose that i_1's true preferences $\succ_{i_1}$ are $s_2 \succ_{i_1} s_\emptyset \succ_{i_1} s_1$. Also, let $\succ_{-i} \in \mathcal{P}_{-i}$ satisfy $s_1 \succ_{i_2} s_2 \succ_{i_2} s_\emptyset$ and $s_2 \succ_{i_3} s_1 \succ_{i_3} s_\emptyset$. Next, consider score structure g with $g^{s_1}_{i_1} > g^{s_1}_{i_3} > g^{s_1}_{i_2}$ and $g^{s_2}_{i_2} > g^{s_2}_{i_1} > g^{s_2}_{i_3}$. Note that the unique stable matching with respect to $\succ$ is $\nu = \{\boldsymbol{(i_1, s_\emptyset)}, (i_2, s_2), (i_3, s_1)\}$ and that matching $\mu = \{\boldsymbol{(i_1, s_\emptyset)}, (i_2, s_1), (i_3, s_2)\}$ is the unique Pareto efficient matching that Pareto dominates ν. Thus, for an arbitrary efficient stable dominating rule, denoted by f^{ESD}, we must have $f^{ESD}(\succ) = \mu$.

In the following, we construct a misreport $\tilde{\succ}_{i_1}$ through which i_1 regrets $\succ_{i_1}$ at observation $(\mu, \pi(\mu, g))$. Before we can make this misreport explicit, we need to describe i_1's inference set $\mathcal{I}(\mu, \succ_{i_1})$. To start, note that $g^{s_1}_{i_1} > \pi_{s_1}(\mu, g)$ and $g^{s_2}_{i_1} > \pi_{s_2}(\mu, g)$. We now show that any $\tilde{g^{s_2}}$ must share its ordinal ranking with g^{s_2} for any plausible score structure $\tilde{g}_{-i}$. First, from the observation $(\mu, \pi(\mu, g))$ student i_1 observes that her top choice s_2 is assigned to a lower priority student i_3, i.e. $\tilde{g}^{s_2}_{i_1} > \tilde{g}^{s_2}_{i_3}$. Second, if i_1 would have top priority at s_2 this would imply that i_1 is assigned to s_2 under any stable matching ν' whenever s_2 is submitted as her top choice. Thus, this must also hold true for any Pareto Efficient matching μ' that improves on ν' and hence i_1 can infer that student i_2 must have top priority at s_2. In conclusion, for any plausible $(\tilde{g}_{-i_1}, \tilde{\succ}_{-i_1})$, the corresponding $\tilde{g}^{s_2}$ shares the same ordinal ranking with g^{s_2}.

Next, given $\tilde{g}^{s_2}$, it must hold $\tilde{\succ}_{i_2} = \succ_{i_2}$. First, i_2 must submit s_2 as acceptable since otherwise any stable matching would assign s_2 to i_1. Therefore, i_1 knows $s_2 \tilde{\succ}_{i_2} s_\emptyset$. Second, note that since i_2 has top priority at s_2, f^{ESD} would have assigned s_2 to i_2 if i_2 would have submitted s_2 as her top choice. Thus, i_1 knows $s_1 \tilde{\succ}_{i_2} s_2$. Combining the two relations i_1 can infer that $\tilde{\succ}_{i_2} = \succ_{i_2}$ is the unique candidate contained in any plausible $(\tilde{g}_{-i_1}, \tilde{\succ}_{-i_1})$.

Now, we describe the candidates for $\tilde{g}^{s_1}$. First, by observing $(\mu, \pi(\mu, g))$, student i_1 knows that s_1 is assigned to the lower priority student i_2, i.e., $\tilde{g}^{s_1}_{i_1} > \tilde{g}^{s_1}_{i_2}$. Second, we establish that given the information regarding $\tilde{g}^{s_2}$ and $\tilde{\succ}_{i_2}$, we must have $\tilde{g}^{s_1}_{i_3} > \tilde{g}^{s_1}_{i_2}$. Suppose by contradiction that $\tilde{g}^{s_1}_{i_2} > \tilde{g}^{s_1}_{i_3}$. In this case, in f^{ESD}, i_1 and i_2 must be assigned to their top choices s_2 and s_1, respectively. However, this is incompatible with μ. Thus, there are two remaining ordinal rankings $\tilde{g}^{s_1}_{i_1} > \tilde{g}^{s_1}_{i_3} > \tilde{g}^{s_1}_{i_2}$ and $\tilde{g}^{s_1}_{i_3} > \tilde{g}^{s_1}_{i_1} > \tilde{g}^{s_1}_{i_2}$ that are compatible with any plausible scenario $(\tilde{g}_{-i_1}, \tilde{\succ}_{-i_1})$.

We show that only $\tilde{\succ}_{i_3} = \succ_{i_3}$ is compatible with i_1's observation. First, since i_3 is assigned to s_2 in μ, student i_1 can conclude that $s_2 \tilde{\succ}_{i_3} s_\emptyset$. If i_3 would have submitted $s_\emptyset \tilde{\succ}_{i_3} s_1$, then any stable matching would have assigned both i_1 and i_2 to their top choices, which is incompatible with the observation. Thus, it must be true that $s_1 \tilde{\succ}_{i_3} s_\emptyset$. Furthermore, suppose by contradiction that $s_1 \tilde{\succ}_{i_3} s_2$. Given that $s_\emptyset \succ_{i_1} s_1$ and $\tilde{g}^{s_1}_{i_3} > \tilde{g}^{s_1}_{i_2}$, student i_3 is assigned to s_1 under f^{ESD}, which is again incompatible with observing μ. Hence, student i_3 can only have submitted $\tilde{\succ}_{i_3} = \succ_{i_3}$. As a result, we can classify i_1's inference set $\mathcal{I}(\mu, \succ_{i_1})$ into two cases that are distinguished by the remaining candidates of ordinal rankings for scores at s_1.

We now show that i_1 regrets reporting the truth $\succ_{i_1}$ at $(\mu, \pi(\mu, g))$ through $\tilde{\succ}_{i_1} : s_2 \tilde{\succ}_{i_1} \boldsymbol{s_1} \tilde{\succ}_{i_1} \boldsymbol{s_\emptyset}$. We show that among the two possible classes from the inference set, in one class i_1 is strictly better off through the misreport and she is not worse off in the remaining class.

First, suppose that $(\tilde{g}_{-i_1}, \tilde{\succ}_{-i_1}) \in \mathcal{I}(\mu, \succ_{i_1})$ satisfies $\tilde{g}^{s_1}_{i_1} > \tilde{g}^{s_1}_{i_3} > \tilde{g}^{s_1}_{i_2}$. In this case, we argue that f^{ESD} must assign i_1 to s_2 when i_1 submits $\tilde{\succ}_i$. Hence, student i_1 would strictly improve her assignment from $s_\emptyset$ under truth-telling to her top choice s_2.

We first show that there is a unique stable matching $\tilde{\nu} = \{\boldsymbol{(i_1, s_1)}, (i_2, s_2), (i_3, s_\emptyset)\}$. Note that in any stable matching i_1 cannot be assigned to $s_\emptyset$, since i_1 would have justified envy at s_1. This implies that whenever i_1 is not assigned to s_2, she must be assigned to s_1. Furthermore, if i_1 is matched with s_2, then i_2 must be assigned to s_1, which would mean that i_3 has justified envy at s_1. Thus, the unique stable matching corresponds to $\tilde{\nu}$. Hence, any efficient stable dominating rule selects $\tilde{\mu} = \{\boldsymbol{(i_1, s_2)}, (i_2, s_1), (i_3, s_\emptyset)\}$ since it is the only Pareto efficient matching that dominates $\tilde{\nu}$. Thus, we conclude that conditional on her observation $(\mu, \pi(\mu, g))$, in this scenario, i_1 would have been better off if she had reported $\tilde{\succ}_{i_1}$ to f^{ESD}.

It remains to show that given $(\tilde{g}_{-i_1}, \tilde{\succ}_{-i_1}) \in \mathcal{I}(\mu, \succ_{i_1})$ with $\tilde{g}^{s_1}_{i_3} > \tilde{g}^{s_1}_{i_1} > \tilde{g}^{s_1}_{i_2}$, student i_1 is not assigned to a worse option than under truth-telling (namely s_1). Clearly, in this case the unique stable matching is ν, while the unique matching that Pareto dominates ν is μ. Therefore, i_1 will be assigned to $s_\emptyset$ under f^{ESD}, which is the same assignment as under true preferences.

Since the choice of f^{ESD} was arbitrary, we have shown that for any efficient stable dominating rule, student i_1 regrets reporting the truth $\succ_{i_1}$ through misreport $\tilde{\succ}_{i_1}$ at $(\mu, \pi(\mu, g))$. This completes the proof. □

As shown next, Theorem 2 cannot be generalized to hold for all stable dominating rules without consent decisions and which are different from DA.

Proposition 3. *There exists a non-stable and non-efficient stable dominating rule without consent decisions which is regret-free truth-telling.*

Proof. See Appendix E. □

The rule we construct in the proof of Proposition 3 always selects the DA outcome except when the input is the same as under the problem studied in the proof of Theorem 2, where it selects an unstable but efficient matching.

As a final remark, note that not all non-stable and non-efficient stable dominating rules are regret-free truth-telling. An example is a modification of the efficient stable dominating DA+TTC, which first runs DA, then gives each student her matched school as an endowment and runs the *Top Trading Cycles (TTC)* algorithm by Shapley and Scarf (1974). More precisely, consider a non-efficient variant of DA+TTC where only cycles that contain exactly two students are executed. A brief inspection of the proof of Theorem 2 shows that this variant of DA+TTC coincides with an efficient stable dominating rule in the relevant case and the proof can be applied directly.

6 Discussion

As we remarked in Section 4, a necessary condition for EDA to be regret-free truth-telling (Theorem 1) is that students face uncertainty regarding the consent decisions of other students. Yet specifying the consent decisions of students is rarely critical to our arguments. With one exception that is discussed below, the consent decisions of all students could be disclosed without affecting the conclusion.

The exception occurs under generalizations of the problem that appears in the proof of Theorem 2. If one would use EDA for this problem with $c_{i_1} = 1$, then i_1 would observe $(\mu, \pi(\mu, g))$. In this case, the observation reveals that i_1 has justified envy at the two schools s_1 and s_2 and i_1 can thus infer that their assigned students must have benefited compared to DA through her own consent. The details i_1 can infer from the observations' features in this example are then rich enough for her to conclude that by misreporting $\tilde{\succ}_{i_1}$, her applications either would have made i_3 a lastly rejected interrupter at school s_2 or that i_1 remains to be matched with her observed matching $s_\emptyset$. In the former case, having $c_{i_3} = 1$ would ensure that i_1 improves from s_1 to s_2 under EDA's updating procedure, whereas $c_{i_3} = 0$ implies that i_1 is matched with her truly least preferred school s_1. Consequently, the truth remains regret-free since no inferences about the consent decisions of i_3 can be drawn from i_1's observation.

We now continue with a short discussion of how robust our main results are to differences in the information structure and our modeling decisions. First, note that if students obtain less detailed information through their observations or if students' choice sets expand, also the set of plausible scenarios weakly expands. Intuitively, in the proof of Theorem 1 we aim at constructing plausible scenarios where students would be harmed by their misreports in a plausible scenario. Thus, with weakly expanding sets of plausible scenarios our arguments would still be valid. For instance, Theorem 1 continues to hold if students only observe their own assignments and the cutoffs of the schools they applied to, or if students can choose their consent decision as a function of the school as in Dur et al. (2019).

For similar reasons, one can extend Theorem 1 to some straightforward environments where scores initially may contain ties between students—a common feature in many applications of school choice. Specifically, given a score structure that contains ties, assume that students do not receive information on the indifference (equivalence) classes of schools' scores, i.e., no student knows at any given school whether another student has the same score as she has. Also assume that students learn their own scores not before the ties have been randomly resolved by a tie-breaker (e.g., by using single tie-breaking or multiple tie-breaking) and that also the observed cutoffs are based on the scores after tie-breaking.[8] It is clear then that for each student, each observation and each strict score structure resulting from tie-breaking that can be part of a plausible scenario, one can separately apply our arguments in the proof of Theorem 1 and thus reach the desired conclusion that truth-telling is regret-free.

We close the discussion with a final note on our negative result Theorem 2. Concretely, consider the case where each student observes only her own assignment and the cutoffs of the schools she applied to. Carefully inspecting the particular problem in the proof of Theorem 2 again, reveals that student i_1 has only one additional

[8]Under a multiple tie-breaking rule, a different tiebreaker can be used at each school, while with a single tie-breaking rule, it is the same for all schools. For a formal treatment of different tie-breaking rules see, for instance, Abdulkadiroğlu et al. (2009).

consistent matching and one additional plausible score ranking for school s_1. In this case, switching the assignments for student i_2 and i_3 compared to μ and using a symmetric argument will lead to the same conclusions as for the original setting.

7 Conclusion

Telling the truth is a safe choice under EDA if students wish to avoid regret their submitted reports. Strengthening this first result, we have also shown that truth-telling is the unique regret-free option under EDA. Moreover, we established that in the class of stable dominating rules without consent decisions, there are candidates which are regret-free truth-telling, whereas no such candidate can be efficient.

Our results open up several avenues for future research. For example, a natural step seems to be to further explore the scope of relaxations of observational constraints that do not affect our results. In another direction, it is also an open question whether EDA is still regret-free if schools' priorities take the form of more flexible choice functions.[9]

Appendix A DA and TP rule

We first introduce the algorithm which induces DA. Thereafter, we present a lemma on DA that is necessary to prove Proposition 1 and Theorem 1 and introduce the TP algorithm. First, fix a problem $(I, S, q, g, \succ, c)$ and consider the DA algorithm:

Step k Each student applies to her most-preferred school $s \in S \cup \{s_\emptyset\}$ that has not rejected her. Each school s tentatively accepts the q_s highest scored students among those who have applied to it (or each of them, if fewer than q_s apply), and rejects the rest.

[9] Ehlers and Morrill (2020) introduce a generalized version of EDA that might serve as a starting point for an investigation.

The algorithm terminates with the tentative assignments of the first step in which no student is rejected. For our lemma presented below we define Weak Maskin Monotonicity as in Kojima and Manea (2010). We call $\succ'$ a monotonic transformation of $\succ$ at matching μ, if for each $i' \in I$, $\succ'_{i'}$ is a monotonic transformation of $\succ_{i'}$ at $\mu_{i'}$.

Definition 4. A matching rule f is *weakly Maskin monotonic* if, given any $\succ$ and for any $\succ'$ that is a monotonic transformation of $\succ$ at $f(g, \succ, c)$, $f(g, \succ', c)$ weakly Pareto dominates $f(g, \succ, c)$

Kojima and Manea (2010) show that DA is weakly Maskin monotonic. Furthermore, DA is strategy-proof (Dubins and Freedman (1981) and Roth (1982)) and produces the SOSM for a given score structure and preference profile.

Lemma 1. *Suppose that $\succ'_i \in \mathcal{P}$ is a monotonic transformation of $\succ_i$ at $DA_i(g, \succ)$. Then, $DA(g, (\succ'_i, \succ_{-i}))$ weakly Pareto dominates $DA(g, \succ)$ and i's outcomes are identical, i.e., $DA_i(g, \succ) = DA_i(g, (\succ'_i, \succ_{-i}))$.*

Proof. The first part follows from weak Maskin monotonicity of DA. The second part is proved by means of contradiction. Suppose that $DA_i(g, \succ) \neq DA_i(g, (\succ'_i, \succ_{-i}))$, then by weak Maskin monotonicity of DA, $DA_i(g, (\succ'_i, \succ_{-i})) \succ_i DA_i(g, \succ)$, which contradicts strategy-proofness of DA. □

Relevant for our proofs, we now introduce how the *Top-Priority (TP) algorithm* (Dur et al., 2019) calculates the outcomes of EDA and start with some basic terminologies. Fix any $(\succ, c)$. For any matching $\mu \in \mathcal{M}$, any student i and any school s, we say that i *demands* s *at* μ if $s \succ_i \mu_i$. Moreover, we say that student i is *eligible* for s at μ if i demands s at μ and there exists no j who also demands s with $c_j = 0$ and $g_i^s < g_j^s$. Note that there could be more than one student who is eligible for a school and if two students i, i' are both eligible for s, then $g_i^s > g_{i'}^s$ implies $c_i = 1$.

Given a matching $\mu \in \mathcal{M}$, consider the directed graph $G(\mu) = (I, E(\mu))$, where $E(\mu) \subseteq I \times I$ is the set of (directed) edges such that $ij \in E(\mu)$ if and only if i is eligible for μ_j. Hence for each student $i \in I$, her directed edges under $G(\mu)$ describe her demands of which the realization would not imply a priority violation given that each student $j \neq i$ is matched with a school weakly preferred to μ_j. A set of edges $\{i_1 i_2, i_2 i_3, ..., i_n i_{n+1}\}$ in $G(\mu)$ is a path if $i_1, i_2, ..., i_{n+1}$ are distinct and it is a cycle if $i_1, i_2, ..., i_n$ are distinct while $i_1 = i_{n+1}$.

A school s has *no demand* in μ if no student demands s at μ. A school s is *underdemanded* at μ if it either has no demand at μ or, every path in $G(\mu)$ that is not part of another path in $G(\mu)$ and that ends with some $i \in \mu_s$, begins with a student assigned to a school with no demand. We say that a student is *permanently matched* at μ if she is assigned to an underdemanded school at μ. Furthermore, a student is *temporarily matched* if she is not permanently matched.

Given $\mu \in \mathcal{M}$, we call $G^*(\mu) = (I, E^*(\mu))$ the *Top-priority graph* of μ and its set of edges $E^*(\mu)$ is defined as follows: we have $ij \in E^*(\mu)$ if and only if among the students who are temporarily matched at μ and are eligible for μ_j, student i has the highest score for μ_j. That is, for each $i \in I$, $E^*(\mu) \subseteq E(\mu)$ contains at most one edge pointing to i. Solving cycle $\gamma = \{i_1 i_2, i_2 i_3, ... i_n i_1\}$ in $G^*(\mu)$ is defined by the operation $\circ$ and yields matching $\nu = \gamma \circ \mu$, such that $\nu_i = \mu_j$ for each $ij \in \gamma$, and $\nu_{i'} = \mu_{i'}$ for each $i' \notin \{i_1, i_2, ..., i_n\}$. The TP algorithm iteratively solves cycles from top-priority graphs as follows:

Step 0**:** Run DA and denote the matching outcome by μ^0.

Step t**:** Given matching μ^{t-1}:

- $t.1$ If there is no cycle in $G^*(\mu^{t-1})$, then stop and let the outcome be μ^{t-1}.
- $t.2$ Otherwise, select one of the cycles in $G^*(\mu^{t-1})$, say γ^t, and let $\mu^t = \gamma^t \circ \mu^{t-1}$. Move to step $t + 1$.

As has been shown in Lemma 6 of Dur et al. (2019), any cycle selection of the algorithm leads to the outcome of EDA and thus the TP algorithm induces EDA.

Appendix B Proof of Proposition 1

In this section, we provide an important lemma to prove Proposition 1 and which is also used in the proof of Theorem 1. We use $EDA(\succ)$ to refer to $EDA(g, (\succ_i, \succ_{-i}), c)$; and $EDA(\tilde{\succ})$ to refer to $EDA(g, (\tilde{\succ}_i, \succ_{-i}), c)$. In a similar way, we use $DA(\succ)$ for $DA(g, (\succ_i, \succ_{-i}))$ and $DA(\tilde{\succ})$ to refer to $DA(g, (\tilde{\succ}_i, \succ_{-i}))$.

Let $pTP^{\succ} = \{\gamma^t\}_{t=1}^T$ be an arbitrary process of the TP algorithm with input $(g, \succ, c)$ that is captured by the series of solved top priority cycles $\{\gamma^t\}_{t=1}^T$. Let $EDA^t(\succ)$ be the outcome of the t_{th} step in $pTP^{\succ}$. Specifically, for each $t \leq T$, γ^t is solved at step t of $pTP^{\succ}$ and we set $EDA^t(\succ) = \gamma^t \circ EDA^{t-1}(\succ)$ with $EDA^0(\succ) = DA(\succ)$. Collect the set of schools to which i is (temporarily) assigned during $pTP^{\succ}$ in $S_i = \{\hat{s} \in S \mid \exists t \in \mathbb{N} : EDA_i^t(\succ) = \hat{s}\}$.

Lemma 2. *If $SU_{\hat{s}}^{\tilde{\succ}_i} \subseteq SU_{\hat{s}}^{\succ_i}$ for all $\hat{s} \in S_i$, then $EDA_i(g, (\tilde{\succ}_i, \succ_{-i}), c)) = \mu_i$.*

Note that the condition in Lemma 2 is satisfied if $\tilde{\succ}_i \in \mathcal{P}$ is such that for all $s, s' \in L_{\mu_i}^{\succ_i}$, $s \succ_i s'$ only if $s \tilde{\succ}_i s'$. Thus, Lemma 2 implies Proposition 1.

Proof. We first prove that $EDA(\succ) = EDA(\tilde{\succ})$ when $c_i = 1$. At the end of the proof we consider the case where $c_i = 0$, for which we establish that $EDA_i(\succ) = EDA_i(\tilde{\succ})$.

Since the outcome of the TP algorithm is invariant in the choice of the cycle solved in each round, it suffices to construct one TP process with input $((\tilde{\succ}_i, \succ_{-i}), c, g)$, denoted by $pTP^{\tilde{\succ}}$, that leads to the same outcome as $pTP^{\succ}$. We make use of the algorithm presented next.

Initialize: Let $t = 1$. Also, let $\nu^0(\tilde{\succ}) = DA(\tilde{\succ})$.

Round $t \leq T$: Let $L^t = \{l \in I \mid \nu_l^{t-1}(\tilde{\succ}) \neq EDA_l^{t-1}(\succ)\}$.

- If each $jk \in \gamma^t$ satisfies that $j, k \in L^t$, let $\nu^t(\tilde{\succ}) = \nu^{t-1}(\tilde{\succ})$. Then, move to Round $t + 1$ or terminate the algorithm if $t = T$.

- If there exists $jk \in \gamma^t$ such that $j \notin L^t$ or $k \notin L^t$, let $\nu^t(\tilde{\succ}) = \gamma^t \circ \nu^{t-1}(\tilde{\succ})$. Then, move to Round $t+1$ or terminate the algorithm if $t = T$.

Collect in $\{\tilde{\gamma}^t\}_{t=1}^{\tilde{T}}$ the series of cycles solved while running the algorithm. By construction, we have $\{\tilde{\gamma}^t\}_{t=1}^{\tilde{T}} \subseteq \{\gamma^t\}_{t=1}^{T}$. We now show that the generated cycle selection $\{\tilde{\gamma}^t\}_{t=1}^{\tilde{T}}$ allows to describe the desired $pTP^{\tilde{\succ}}$. Our strategy will be as follows. We establish in the first step that the algorithm is well defined. In the second step, we will argue that $\nu^T(\tilde{\succ}) = EDA^T(\succ)$ and that $G^*(\nu^T(\tilde{\succ}))$ contains no cycles.

Step 1 We can generate the desired sequence of cycles $\{\tilde{\gamma}^t\}_{t=1}^{\tilde{T}}$ if for each round $t \leq T$, the following four statements are satisfied:

(B1) Either all students involved in γ^t belong to L^t, or none of them does.

(B2) $\gamma^t \in G^*(\nu^{t-1}(\tilde{\succ}))$ when γ^t contains no student from L^t.

(B3) $\nu^t(\tilde{\succ})$ weakly Pareto dominates $EDA^t(\succ)$, and $L^{t+1} \subseteq L^t$.

(B4) For each $l \in L^t$, $DA_l(\tilde{\succ}) = \nu_l^{t-1}(\tilde{\succ})$.

We prove by means of induction that $(B1)$ - $(B4)$ hold for each round of the process. Since the arguments for the intial step and the inductive step are similar and to avoid lengthy repetition of arguments, we establish $(B1)$ - $(B4)$ to be applicable for both the initial step and the inductive step. That is, to apply the arguments for round 1, set $t = 1$ and for $t > 1$, use the inductive hypothesis that $(B1)$-$(B4)$ hold for all rounds $t' < t$.

More specifically, given the induction hypothesis, for each t, statement $(B1)$ is needed to ensure that statement $(B2)$ is true. We then use $(B1)$ and $(B2)$ to establish $(B3)$ and then establish $(B4)$.

For the initial case we build on the following observations. We have $DA_i(\succ) \in S_i$. Hence $\tilde{\succ}_i$ is a monotonic transformation of $\succ_i$ at $DA_i(\succ)$. It is then immediate from

Lemma 1 that $DA(\tilde{\succ})$ weakly Pareto dominates $DA(\succ)$ and $DA_i(\succ) = DA_i(\tilde{\succ})$. Thus, $L^1 = \{l \in I \mid DA_l(\tilde{\succ}) \succ_l DA_l(\succ)\}$ and $i \notin L^1$. Furthermore, by definition it is true that $DA_l(\tilde{\succ}) = \nu_l^0(\tilde{\succ})$ for any $l \in I$. Moreover, let $S' = \{s \in S \mid s \succ_i \mu_i \text{ and } \mu_i \tilde{\succ}_i s\}$.

Statement (B1): Since γ^t is a cycle, it suffices to show that for each $jk \in \gamma^t$, $k \in L^t$ implies $j \in L^t$. We first establish that for any $jk \in \gamma^t$, if $k \in L^t$, then either (1) $j \in L^t$ or (2) $j = i$ and $EDA_k^{t-1}(\succ) \in S'$. By contradiction, let $k \in L^t$, $j \notin L^t$ and if $j = i$, then $EDA_k^{t-1}(\succ) \notin S'$. We aim at a contradiction towards the stability of $DA(\tilde{\succ})$. First, if $k \in L^t$, then there exists $l \in L^t$ such that $\nu_l^{t-1}(\tilde{\succ}) = EDA_k^{t-1}(\succ)$. Now, since $l \in L^t$, it must be true that $DA_l(\tilde{\succ}) = \nu_l^{t-1}(\tilde{\succ}) \succ_l EDA_l^{t-1}(\succ)$. For the initial case this argument is immediate since $DA_l(\tilde{\succ}) = \nu_l^0(\tilde{\succ}) \succ_l DA_l(\succ)$. For $t > 1$, the relation is consequence of the inductive hypothesis. Specifically, $(B4)$ holding in all previous rounds establishes the left side of the relation, and $(B3)$ holding for all previous rounds implies the right side of the relation. Next, together with $jk \in G^*(EDA^{t-1}(\succ))$ this implies that $g_j^{DA_l(\tilde{\succ})} > g_l^{DA_l(\tilde{\succ})}$ and $EDA_k^{t-1}(\succ) \succ_j EDA_j^{t-1}(\succ)$. Furthermore, $j \notin L^t$ implies $EDA_j^{t-1}(\succ) = \nu_j^{t-1}(\tilde{\succ}) \succeq_j DA_j(\tilde{\succ})$. In the following, let $\succ_j = \tilde{\succ}_j$ if $j \neq i$. Now note that $\succ_j = \tilde{\succ}_j$ implies that $EDA_k^{t-1}(\succ) \tilde{\succ}_j EDA_j^{t-1}(\succ)$. Similarly, for $i = j$, if $EDA_k^{t-1}(\succ) \notin S'$, then since for all $\hat{s} \in S_i$, $SU_{\hat{s}}^{\tilde{\succ}_i} \subseteq SU_{\hat{s}}^{\succ_i}$ the variations on $\tilde{\succ}_i$ relative to $\succ_i$ cannot change the position of $EDA_i^{t-1}(\succ)$ relative to $EDA_k^{t-1}(\succ)$ and thus $EDA_k^{t-1}(\succ) \tilde{\succ}_i EDA_i^{t-1}(\succ)$. Thus, combining the relations derived so far means for each $j \notin L^t$ that

$$DA_l(\tilde{\succ}) = \nu_l^{t-1}(\tilde{\succ}) = EDA_k^{t-1}(\succ) \tilde{\succ}_j EDA_j^{t-1}(\succ) = \nu_j^{t-1}(\tilde{\succ}) \tilde{\succeq}_j DA_j(\tilde{\succ})$$

However, this implies that j has justified envy towards l at $DA(\tilde{\succ})$. Hence we arrive at a contradiction to the stability of $DA(\tilde{\succ})$ with respect to $\tilde{\succ}$.

Note that the statement we established above implies that for any $jk \in \gamma^t$, if $k \in L^t$ and $j \neq i$, we have $j \in L^t$. Moreover, the arguments we used ensure that the implication would hold more generally, i.e., for any $jk \in G^*(EDA^{t-1}(\succ))$, if $k \in L^t$ and $j \neq i$, we have $j \in L^t$. This generalization will turn out to be useful in the upcoming arguments in Step 2.

We next show that $jk \in \gamma^t$ and $k \in L^t$ imply $j \neq i$. Based on the statement already established, it suffices to show that $j = i$ and $EDA_k^{t-1}(\succ) \in S'$ is impossible. If $ik \in \gamma^t$ and $EDA_k^{t-1}(\succ) \in S'$, then it implies that $EDA_k^{t-1}(\succ) = EDA_i^t(\succ) \succ_i \mu_i$. However, this is a contradiction to μ being the final matching of $pTP^{\succ}$. Thus, we must have $j \neq i$.

We conclude that once there is an edge $jk \in \gamma^t$ with $k \in L^t$, then $j \in L^t$. Therefore, either all students involved in γ^t belong to L^t, or no such student does.

Statement (B2): Given that (B1) is true at round t, we proceed to prove (B2). Suppose that for each $jk \in \gamma^t$, $j, k \notin L^t$. Thus, we get $EDA_j^{t-1}(\succ) = \nu_j^{t-1}(\tilde{\succ})$ and $EDA_k^{t-1}(\succ) = \nu_k^{t-1}(\tilde{\succ})$. This implies that $\nu_k^{t-1}(\tilde{\succ}) \tilde{\succ}_j \nu_j^{t-1}(\tilde{\succ})$. Note that this also holds if $j = i$, since $EDA_k^{t-1}(\succ) \notin S'$ implies that variations on $\tilde{\succ}_j$ relative $\succ_j$ cannot change the position of $EDA_j^{t-1}(\succ)$ relative to $EDA_k^{t-1}(\succ)$ and thus $EDA_k^{t-1}(\succ) \tilde{\succ}_j EDA_j^{t-1}(\succ)$. Hence, we obtain that student j must still desire $\nu_k^{t-1}(\tilde{\succ})$ at $\nu^{t-1}(\tilde{\succ})$. Clearly, the last argument is true for all j such that $jk \in \gamma^t$. Thus, we have that all students involved in γ^t are temporarily matched at $\nu^{t-1}(\tilde{\succ})$. Next, since $\nu^{t-1}(\tilde{\succ})$ weakly Pareto dominates $EDA^{t-1}(\succ)$, there are weakly fewer temporarily matched students who desire $\nu_k^{t-1}(\tilde{\succ})$ at $\nu^{t-1}(\tilde{\succ})$ compared to $EDA^{t-1}(\succ)$. As a result, j still has the highest score among all temporarily matched students pointing to k. Hence $jk \in G^*(\nu^{t-1}(\tilde{\succ}))$. Since this holds for all edges in γ^t, it follows that $\gamma^t \in G^*(\nu^{t-1}(\tilde{\succ}))$.

Statement (B3): We first show that $\nu^t(\tilde{\succ})$ weakly Pareto dominates $EDA^t(\succ)$. Note that $\nu^{t-1}(\tilde{\succ})$ weakly Pareto dominates $EDA^{t-1}(\succ)$. At $t = 1$ this follows from Lemma 1 and in any round $t > 1$ it follows from the induction hypothesis. Moreover, only students in γ^t change their assignments in round t of our algorithm (and also in round t of $pTP^{\succ}$). Thus, to conclude that $\nu^t(\tilde{\succ})$ weakly Pareto dominates $EDA^t(\succ)$, it is sufficient to show that for each $jk \in \gamma^t$ it holds $\nu_j^t(\tilde{\succ}) \succeq_j EDA_j^t(\succ)$.

Of the two cases we have to consider, we start with the simpler one, in which for any $jk \in \gamma^t$, we have $j, k \notin L^t$. In this case, γ^t is solved in both $\nu^{t-1}(\tilde{\succ})$ and $EDA^{t-1}(\succ)$. Therefore, $\nu_j^t(\tilde{\succ}) = EDA_j^t(\succ)$ and we obtain the desired result.

In the remaining case, any $jk \in \gamma^t$ satisfies that $j, k \in L^t$. Clearly, we can solve a cycle of this form only if $L^t \neq \emptyset$. Moreover, note that $EDA^t(\succ) = \gamma^t \circ EDA^{t-1}(\succ)$ and $\nu^t(\tilde{\succ}) = \nu^{t-1}(\tilde{\succ})$. We proceed by contradiction and assume that $EDA_j^t(\succ) \succ_j \nu_j^t(\tilde{\succ})$. Similar as in the arguments of $(B1)$, we will contradict the stability of $DA(\tilde{\succ})$. We make the following observations: First, since we have $k \in L^t$, there must exist $l \in L^t$ such that we have $\nu_l^{t-1}(\tilde{\succ}) = EDA_k^{t-1}(\succ)$. Second, note that $l \in L^t$ implies the relation $DA_l(\tilde{\succ}) = \nu_l^{t-1}(\tilde{\succ}) \succ_l EDA_l^{t-1}(\succ)$. Therefore, $jk \in \gamma^t$ also means that $g_j^{DA_l(\tilde{\succ})} > g_l^{DA_l(\tilde{\succ})}$ and $EDA_k^{t-1}(\succ) = EDA_j^t(\succ)$. Third, the algorithm guarantees that $\nu_j^t(\tilde{\succ}) \succeq_j DA_j(\tilde{\succ})$. If we combine all relations above with $\succ_j = \tilde{\succ}_j$, we obtain:

$$DA_l(\tilde{\succ}) = \nu_l^{t-1}(\tilde{\succ}) = EDA_k^{t-1}(\succ) = EDA_j^t(\succ) \tilde{\succ}_j \nu_j^t(\tilde{\succ}) \tilde{\succeq}_j DA_j(\tilde{\succ})$$

and reach a contradiction, since j has justified envy towards l at $DA(\tilde{\succ})$. Thus, $\nu^t(\tilde{\succ})$ weakly Pareto dominates $EDA^t(\succ)$. Moreover, based on the weak Pareto dominance we just established, we can write L^{t+1} as $L^{t+1} = \{l \in I \mid \nu_l^t(\tilde{\succ}) \succ_l EDA_l^t(\succ)\}$.

To finish the proof for statement (B3) we need to show that $L^{t+1} \subseteq L^t$. If any $jk \in \gamma^t$ satisfies $j, k \notin L^t$, then it is immediate that $L^{t+1} = L^t$. On the contrary, if any $jk \in \gamma^t$ satisfies $j, k \in L^t$, then, first, for each such j, as $j \in L^t$, we have $\nu_j^{t-1}(\tilde{\succ}) \succ_j EDA_j^{t-1}(\succ)$ and $\nu_j^t(\tilde{\succ}) \succeq_j EDA_j^t(\succ)$. This implies that while j is contained in L^t, she might not be in L^{t+1}. Second, for each $j' \in I$ not involved in γ^t, we have $\nu_{j'}^t(\tilde{\succ}) = \nu_{j'}^{t-1}(\tilde{\succ})$ and $EDA_{j'}^t(\succ) = EDA_{j'}^{t-1}(\succ)$, which implies that $j' \in L^t$ if and only if $j' \in L^{t+1}$. In conclusion, we can infer that $L^{t+1} \subseteq L^t$. Hence (B3) is satisfied.

Statement (B4): For $t = 1$, the statement is immediate. Let $t > 1$. By the inductive hypothesis (in particular ($B3$)), it holds $L^{t'+1} \subseteq L^{t'}$ for any $t' < t$. This implies that $L^t \subseteq L^{t'}$. Second, solving the cycles in the algorithm under the inductive hypothesis implies that, given any $t' < t$, the assignments at $\nu^{t'}(\tilde{\succsim})$ and $\nu^{t'-1}(\tilde{\succsim})$ are identical for each student in $L^{t'}$. Thus, since $L^t \subseteq L^{t'}$, we can infer that for each $l \in L^t$, $DA_l(\tilde{\succsim}) = \nu_l^{t-1}(\tilde{\succsim})$.

Step 2: We show that $EDA^T(\succ) = \nu^T(\tilde{\succsim})$. Let $t_i \leq T$ be the first step in $pTP^{\succ}$ where i is permanently matched and consider round t_i of our algorithm. If $EDA^{t_i-1}(\succ) = \nu^{t_i-1}(\tilde{\succsim})$, we have that $L^t = \emptyset$ and that γ^t is solved in each round $t > t_i$ of the algorithm. Consequently, it is true that $EDA^T(\succ) = \nu^T(\tilde{\succsim})$. If $EDA^{t_i-1}(\succ) \neq \nu^{t_i-1}(\tilde{\succsim})$, then L^{t_i} is non-empty. In this case, we show that there exists $\hat{t} > t_i$ such that $EDA^{\hat{t}}(\succ) = \nu^{\hat{t}}(\tilde{\succsim})$. As shown above, this leads to $EDA^T(\succ) = \nu^T(\tilde{\succsim})$.

We show that there must be a cycle in $G^*(EDA^{t_i-1}(\succ))$ that solely consists of elements in L^{t_i}. We begin with showing that for any $k \in L^{t_i}$, there exists an edge $jk \in G^*(EDA^{t_i-1}(\succ))$ for some $j \in I$. Since $k \in L^{t_i}$, there exists $l \in L^{t_i}$ such that $EDA_k^{t_i-1}(\succ) = \nu_l^{t_i-1}(\tilde{\succsim}) \succ_l EDA_l^{t_i-1}(\succ)$. That is, at $EDA^{t_i-1}(\succ)$, for each student in L^{t_i}, her assignment is desired by at least one student in L^{t_i} whose assignment is further desired by some other student in L^{t_i}. Now, recall that we assume $c_1 = 1$. Since i is permanently matched at step t_i and i consents, then even if i prefers $EDA_k^{t_i-1}(\succ)$ to μ_i, she cannot prevent any student from being eligible for $EDA_k^{t_i-1}(\succ)$. In other words, at least one edge that is pointing to k, namely lk, is contained in $G(EDA^{t_i-1}(\succ))$. Therefore, we can infer that k is temporarily matched in $EDA^{t_i-1}(\succ)$ and thus there must be $jk \in G^*(EDA^{t_i-1}(\succ))$ for some $j \in I$.

Next, for any such jk, our arguments from (B1) will be sufficient to conclude that $j \in L^{t_i}$. First, we have already shown $j \in L^{t_i} \cup \{i\}$. Second, we know that $j \neq i$, since i is permanently matched. Thus, we can infer that each student in L^{t_i} is pointed by another student in L^{t_i} in $G^*(EDA^{t_i-1}(\succ))$. Since L^{t_i} is finite, the existence of the desired cycle is guaranteed. Notably, according to (B3) and by iteratively applying the same argument, we can eventually reach a round $\hat{t} > t_i$ where $EDA^{\hat{t}}(\succ) = \nu^{\hat{t}}(\tilde{\succ})$.

We next claim that no cycles can be found in $G^*(\nu^T(\tilde{\succ}))$. Notably, if $G^*(\nu^T(\tilde{\succ}))$ has a cycle, then using the arguments in (B2) implies that $G^*(EDA^T(\succ))$ must also have a cycle. However, this contradicts the fact that exactly T cycles are solved in $pTP^{\succ}$. Based on the statements provided so far, we can construct the desired $pTP^{\tilde{\succ}}$ as $pTP^{\tilde{\succ}} = \{\tilde{\gamma}^t\}_{t=1}^{\tilde{T}}$. Thus, $EDA(\succ) = EDA(\tilde{\succ})$ which completes the proof for $c_i = 1$.

Finally, we extend the arguments to the case where $c_i = 0$. Note that EDA is consent-invariant and thus $EDA_i(\succ) = EDA_i(g, (\succ_i, \succ_{-i}), (\tilde{c}_i, c_{-i}))$ and also $EDA_i(\tilde{\succ}) = EDA_i(g, (\tilde{\succ}_i, \succ_{-i}), (\tilde{c}_i, c_{-i}))$ for $\tilde{c}_i = 1$. Moreover, we have just shown that when i consents, submitting $\tilde{\succ}_i$ will not alter the EDA outcome, that is, $EDA(g, (\succ_i, \succ_{-i}), (\tilde{c}_i, c_{-i})) = EDA(g, (\tilde{\succ}_i, \succ_{-i}), (\tilde{c}_i, c_{-i}))$. This allows us to conclude $EDA_i(\succ) = EDA_i(\tilde{\succ})$, which completes the proof. □

Appendix C Proof of Theorem 1

Fix an arbitrary problem $(I, S, q, g, \succ, c)$ and consider an arbitrary student $i \in I$. Since EDA only takes acceptable schools into account, for any tuple $(g, \succ_{-i}, c)$ and any $\succ'_i$ which is truth-telling, we have $EDA(g, (\succ'_i, \succ_{-i}), c) = EDA(g, (\succ_i, \succ_{-i}), c)$. Hence, if student i does not regret reporting her true preferences $\succ_i$, she does not regret to report any truth-telling report $\succ'_i$. Thus, we show that i does not regret to report $\succ_i$.

Lemmas 3, 5 and 9 will each consider a distinct class of misreports of student i and jointly imply that i cannot regret submitting her true preferences. In the following exposition, take an arbitrary observation $(\mu, \pi(\mu, g))$ where $\mu \in \mathcal{M}|_{(\succ_i, c_i)}$. We fix i's scores g_i and i's consent decision c_i throughout the proof. From now on, we use $\tilde{g}$ to refer to $(g_i, \tilde{g}_{-i})$ and $\tilde{c}$ to refer to $(c_i, \tilde{c}_{-i})$.

We first show that a misreport is not profitable for i, if it shares the same relative ranking of schools weakly below her own assignment under truth-telling.

Lemma 3. *Consider* $\tilde{\succ}_i \in \mathcal{P}$ *such that for all* $s, s' \in L^{\succ_i}_{\mu_i}$, $s \tilde{\succ}_i s'$ *if and only if* $s \succ_i s'$. *For any* $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$, *it is true that* $EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = \mu_i$.

Proof. Select any $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$. By definition, $EDA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}), \tilde{c}) = \mu$ and using Proposition 1, we know $EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = \mu_i$. □

Before formally presenting our arguments for other misreports, we provide the following auxiliary result.

Lemma 4. *Fix any* $\hat{\succ} \in \mathcal{P}_I$, *any* $\hat{g} \in \mathcal{G}_I$ *and any* $\hat{c} \in \mathcal{C}_I$. *If* $DA_j(\hat{g}, \hat{\succ}) \hat{\succeq}_j DA_i(\hat{g}, \hat{\succ})$ *for all* $j \in I$, *then* $EDA_i(\hat{g}, \hat{\succ}, \hat{c}) = DA_i(\hat{g}, \hat{\succ})$.

Proof. Note that $DA_j(\hat{g}, \hat{\succ}) \hat{\succeq}_j DA_i(\hat{g}, \hat{\succ})$ for all $j \in I$, implies $DA_j(\hat{g}, \hat{\succ}) \hat{\succ}_j DA_i(\hat{g}, \hat{\succ})$ for any $j \in I$ such that $DA_j(\hat{g}, \hat{\succ}) \neq DA_i(\hat{g}, \hat{\succ})$. That means, $DA_i(\hat{g}, \hat{\succ})$ has no demand at $DA(\hat{g}, \hat{\succ})$. Therefore, $DA_i(\hat{g}, \hat{\succ})$ is underdemanded at $DA(\hat{g}, \hat{\succ})$ and i will not be involved in any cycle solution during any process calculating $EDA(\hat{g}, \hat{\succ}, \hat{c})$. As a result, we have $EDA_i(\hat{g}, \hat{\succ}, \hat{c}) = DA_i(\hat{g}, \hat{\succ})$. □

In the remainder of the proof, the following argument is applied repeatedly for the remaining categories of misreports: When i submits misreport $\tilde{\succ}_i$, then there is a plausible scenario $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$ such that we can apply Lemma 4 under $(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c})$. Moreover, in this case, we will show that $\mu_i \succ_i DA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}))$.

We proceed with misreports in which some schools ranked below μ_i under truth permute their order with μ_i. Our next Lemma shows that the student can either infer that she would have possibly been worse off, or that the misreport would not have affected her assignment in any plausible scenario.

Lemma 5. *Consider* $\tilde{\succ}_i \in \mathcal{P}$ *such that* $\mu_i \succ_i s$ *and* $s \tilde{\succ}_i \mu_i$ *for some* $s \in S$. *Then, either* (1) *there exists* $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$ *such that* $\mu_i \succ_i EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c})$ *or* (2) *for any* $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$: $EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = \mu_i$.

Proof. Let $\tilde{S} = \{s' \in S \mid \mu_i \succ_i s' \text{ and } s' \tilde{\succ}_i \mu_i\}$. We start with a singelton $\tilde{S} = \{s^*\}$ and generalize the arguments later on. We now distinguish the following exhaustive cases based on i's observation $(\mu, \pi(\mu, g))$:

Case 1: $\pi_{s^*}(\mu, g) = 0$. Note that s^* has vacant seat at $EDA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}), \tilde{c}) = \mu$, for any $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$. Thus, at $DA(\tilde{g}(\succ_i, \tilde{\succ}_{-i}))$, s^* must also have a vacant seat and for any $i' \in I$, i' weakly prefers $DA_{i'}(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}))$ to s^* given $\tilde{\succ}_{i'}$. Hence s^* has no demand.

Next, if i submits $\tilde{\succ}_i$, then we obtain $DA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = s^*$. Now notice that before being matched to the final assignment, the set of applications i sends to reach $DA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}))$ is a subset of those sent to reach $DA_i(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}))$. Therefore, each student $i' \neq i$ must weakly prefer $DA_{i'}(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}))$ to $DA_{i'}(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}))$ given her preferences are $\tilde{\succ}_{i'}$. Accordingly, each student $i' \in I$ still weakly prefers $DA_{i'}(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}))$ to s^* given her preferences are $\tilde{\succ}_{i'}$. By Lemma 4, we thus have $EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = DA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = s^*$: Statement (1) holds.

Case 2: $\pi_{s^*}(\mu, g) \neq 0$**,** $\pi_{\mu_i}(\mu, g) = 0$ **and** $g_i^{s^*} < \pi_{s^*}(\mu, g)$**.** We show that statement (2) is satisfied. Take an arbitrary $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$. To start, note that whenever a student j improves her assignment from one school to another at one step of the TP algorithm, another student with lower score is assigned to the school that j left at that step. Since $g_i^{s^*} < \pi_{s^*}(\mu, g)$, this implies that student i must have a lower score than any student assigned to s^* at $DA_i(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}))$. Thus, compared to the DA procedure of i submitting $\succ_i$, i's additional application to s^* by submitting $\tilde{\succ}_i$ has no influence on the outcome and we reach $DA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = DA(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}))$. Moreover, since $\pi_{\mu_i}(\mu, g) = 0$, non-wastefulness of DA implies that all students weakly prefer their assignments to μ_i at $DA(\tilde{g}, (\tilde{\succ}_i . \tilde{\succ}_{-i}))$. We then apply Lemma 4 and conclude $EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = DA_i(\tilde{g}, (\tilde{\succ}_i . \tilde{\succ}_{-i})) = \mu_i$: Statement (2) holds.

Case 3: $\pi_{s^*}(\mu, g) \neq 0$ **and either (C1)** $g_i^{s^*} > \pi_{s^*}(\mu, g)$**; or (C2)** $\pi_{\mu_i}(\mu, g) \neq 0$ **and** $g_i^{s^*} < \pi_{s^*}(\mu, g)$**.**[10] Except for Case 3.2.2.2, statement (1) will apply and our approach is standardized as follows:

Step 1: We construct a candidate scenario $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i})$.

Step 2: We show that $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$.

Step 3: We argue that $EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = s^*$.

Let $j \in I$ be such that $\mu_j = s^*$ and $g_j^{s^*} = \pi_{s^*}(\mu, g)$. Let $\hat{S} = \{s_1, \dots, s_T\}$ be the set of schools for which i has justified envy at μ and assume without loss of generality that $s_1 \succ_i s_2 \succ_i \dots \succ_i s_T$. For any $\succ_i' \in \mathcal{P}$ and $s \in S$, denote the strict lower contour set of $\succ_i'$ at s by $SL_s^{\succ_i'} = \{s' \in S \mid s \succ_i' s'\}$ and the strict upper contour set of $\succ_i'$ at s by $SU_s^{\succ_i'} = \{s' \in S \mid s' \succ_i' s\}$. The following observations on $\hat{S}$ will be helpful:

- $\hat{S} = \emptyset$, if $c_i = 0$, since EDA does not allow for any priority violations for i.
- Non-wastefulness of EDA implies that for each $s' \in \hat{S}$, $\pi_{s'}(\mu, g) \neq 0$.
- Since $\hat{S} \subseteq SU_{\mu_i}^{\succ_i}$ and $s^* \in SL_{\mu_i}^{\succ_i}$, $s^* \notin \hat{S}$.

Now, for each $t \in \{1, \dots, T\}$, let $i_t \in \mu_{s_t}$ be such that $g_{i_t}^{s_t} = \pi_{s_t}(\mu, g)$. Collect all such students in $\hat{I} = \{i_1, \dots, i_T\}$. Note that for each $i_t \in \hat{I}$, in any TP process corresponding to a plausible scenario, there must exist a solved cycle γ such that $i_t k \in \gamma$ for some $k \in I$ and i_t is assigned to s_t when γ is solved. Moreover, solving γ must be the last step in that TP process in which i_t is improved. We distinguish cases by different cardinalities of $\hat{S}$.

Case 3.1: $|\hat{S}| \neq 1$**.** For now, assume that (C2) is satisfied.

Step 1: We start with the candidate score structure $\tilde{g}_{-i}$:

[10] Since we assume that $g_i^s \neq g_j^s$ for any $i, j \in I$ and any $s \in S$, note that it cannot be true that $\pi_{s^*}(\mu, g) = g_i^{s^*}$, when $i \notin \mu_{s^*}$.

- let $g_i^{\mu_i} \geq \pi_{\mu_i}(\mu, g) > \tilde{g}_j^{\mu_i}$ and let $\tilde{g}_k^{\mu_i} = g_k^{\mu_i}$ for all $k \in I \setminus \{i, j\}$ and;
- for any $s' \in S \setminus \{\hat{S} \cup \mu_i\}$, let $\tilde{g}^{s'} = g^{s'}$.

Let $i_0 = i_T$ and $s_{T+1} = s_1$. In case that $\hat{S} \neq \emptyset$, let for each $s_t \in \hat{S}$, be $\tilde{g}^{s_t}$ such that $\tilde{g}_{i_{t-1}}^{s_t} > g_i^{s_t} > \tilde{g}_{i_t}^{s_t}$ with $\tilde{g}_{i_t}^{s_t} = \pi_{s_t}(\mu, g)$ and for all $l \in \mu_{s_t}$ with $l \neq i_t$, let $\tilde{g}_l^{s_t} > \tilde{g}_{i_{t-1}}^{s_t}$. Next, select an arbitrary $\tilde{c}_{-i}$ and consider the following preferences $\tilde{\succ}_{-i}$:

$$\mu_i \; \tilde{\succ}_j \; s^* \; \tilde{\succ}_j \; s_\emptyset \; \tilde{\succ}_j \; \ldots,$$

$$s_t \; \tilde{\succ}_{i_t} \; s_{t+1} \; \tilde{\succ}_{i_t} \; s_\emptyset \; \tilde{\succ}_{i_t} \; \ldots \quad \forall t \in \{1, \ldots, T\},$$

$$\mu_k \; \tilde{\succ}_k \; s_\emptyset \; \tilde{\succ}_k \; \ldots \quad \forall k \in I \backslash (\hat{I} \cup \{i, j\}),$$

Step 2: The construction of $\tilde{g}_{-i}$ ensures that for each $s \in S \setminus \hat{S}$ and each $k \in \mu_s$, we have $\tilde{g}_k^s = g_k^s$. Also, the construction of $\tilde{g}^{s_t}$ for each $s_t \in \hat{S}$ guarantees that $\pi_{s_t}(\mu, (g_i, \tilde{g}_{-i})) = \tilde{g}_{i_t}^s = \pi_{s_t}(\mu, g)$. Thus, we can infer $\pi(\mu, (g_i, \tilde{g}_{-i})) = \pi(\mu, g)$.

We next show that the constructed scenario $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i})$ yields μ under the TP algorithm. First, if $\hat{S} = \emptyset$, we get $DA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = \mu$ and the TP process terminates with μ since there are no cycles $G^*(\mu)$. Second, suppose that $\hat{S} \neq \emptyset$. We describe how we arrive at the corresponding DA outcome: $DA_k(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = \mu_k$ for all $k \in I \setminus \hat{I}$ and $DA_{i_t}(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = s_{t+1}$ for all $i_t \in \hat{I}$. Each $k \in I \setminus \{i, j\}$ is accepted by her top choice μ_k at step 1. Moreover, at some step, student i applies to s_1 and gets tentatively accepted. For each $t \in \{1, \ldots, T\}$, this leads to i_t getting rejected by s_t and applying to s_{t+1} in the next step, causing i_{t+1} being rejected by s_{t+1} and so forth. Eventually i is rejected by s_1, applies to all schools in $SU_{\mu_i}^{\succ_i} \setminus SU_{s_1}^{\succ_i}$ being finally accepted by μ_i. Thus, j is rejected by μ_i and is accepted by s^*.

Next, there is a unique cycle $\gamma = \{i_T i_{T-1}, \ldots, i_2 i_1, i_1 i_T\}$ in $G^*(DA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})))$ which, once solved, produces μ. According to $(\succ_i, \tilde{\succ}_{-i})$, i and j are the only students who do not receive their top choice in μ and therefore the TP algorithm terminates with μ.

Step 3: Be aware that the outcome $DA(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}))$ may vary in the position of s^* on $\tilde{\succ}_i$: If $s^* \tilde{\succ}_i s_1$, then $DA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = s^*$, $DA_j(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = \mu_i$ and $DA_k(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = \mu_k$ for any $k \in I\backslash\{i,j\}$. If $s_1 \tilde{\succ}_i s^*$, then we have $DA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = s^*$, $DA_j(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = \mu_i$, and $DA_{i_t}(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = s_{t+1}$ for $i_t \in \hat{I}$ and $DA_k(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = \mu_k$ for any $k \in I \setminus (\{i,j\} \cup \hat{S})$.

In both instances above, we can apply Lemma 4 to have $EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = s^*$ and the argument for $(C2)$ is complete.

Now suppose that $(C1)$ holds.

Step 1: Modify the preferences of j to be $s^* \tilde{\succ}_j s_\emptyset \tilde{\succ}_j \ldots$ and keep all other details of our construction the same as in instance (C2) above.

Step 2 and Step 3: The arguments resemble those in instance (C2) above.

Case 3.2: $|\hat{S}| = 1$.

Case 3.2.1: There exists $s' \in S \setminus \{s_1, \mu_i, s^*\}$ such that $\pi_{s'}(\mu, g) \neq 0$. Pick an arbitrary such s' and denote with j' the student who has the lowest score among all students being assigned to s' under μ.

Step 1: Let $\tilde{g}_{-i}$ be such that

- $\tilde{g}^{s_1}_{j'} > g^{s_1}_i > \tilde{g}^{s_1}_{i_1}$ and $\tilde{g}^{s_1}_k = g^{s_1}_k$ for all $k \in I \setminus \{i, j'\}$ and;
- $\tilde{g}^{s'}_{i_1} > \tilde{g}^{s'}_{j'}$ and $\tilde{g}^{s'}_k = g^{s'}_k$ for all $k \in I \setminus \{i_1\}$ and;
- $g^{\mu_i}_i > \tilde{g}^{\mu_i}_j$ and $\tilde{g}^{\mu_i}_k = g^{\mu_i}_k$ for all $k \in I \setminus \{i, j'\}$ and;
- $\tilde{g}^{s''} = g^{s''}$ for any $s'' \in S \setminus \{s_1, \mu_i, s'\}$.

Next, fix an arbitrary $\tilde{c}_{-i}$ and consider the following profile $\tilde{\succ}_{-i}$:

$$\mu_i \tilde{\succ}_j s^* \tilde{\succ}_j s_\emptyset \tilde{\succ}_j \ldots,$$

$$s_1 \tilde{\succ}_{i_1} s' \tilde{\succ}_{i_1} s_\emptyset \tilde{\succ}_{i_1} \ldots,$$

$$s' \tilde{\succ}_{j'} s_1 \tilde{\succ}_{j'} s_\emptyset \tilde{\succ}_{j'} \ldots,$$

$$\mu_k \tilde{\succ}_k s_\emptyset \tilde{\succ}_k \ldots \quad \forall k \in I \backslash \{i, i_1, j, j'\}.$$

Step 2 and Step 3: We omit the arguments for Step 2 and Step 3. They are almost identical to those in Case 3.1 and we can eventually apply Lemma 4.

Case 3.2.2: There does not exist $s' \in S \setminus \{s_1, \mu_i, s^*\}$ such that $\pi_{s'}(\mu, g) \neq 0$. Note that this subcase is very specific, as there are only three schools that exhaust their capacity. Here, we have two more subdivisions to make.

Case 3.2.2.1: $g_i^{s^*} > \pi_{s^*}(\mu, g)$. That is, $(C1)$ holds and we have $g_i^{s^*} > g_j^{s^*}$.

Step 1: Let $\tilde{g}_{-i}$ be such that

- $\tilde{g}_j^{s_1} > g_i^{s_1} > \tilde{g}_{i_1}^{s_1}$ and $\tilde{g}_k^{s_1} = g_k^{s_1}$ for all $k \in I \setminus \{i, j\}$ and;
- $g_i^{s^*} > \tilde{g}_{i_1}^{s^*} > \tilde{g}_j^{s^*}$ and $\tilde{g}_k^{s^*} = g_k^{s^*}$ for all $k \in I \setminus \{i, i_1\}$ and;
- $\tilde{g}^{s'} = g^{s'}$ for any $s' \in S \setminus \{s^*, s_1\}$.

Now, let $\tilde{c}_{-i}$ be such that $\tilde{c}_{i_1} = 0$[11] and consider the following profile $\tilde{\succ}_{-i}$:

$$s^* \tilde{\succ}_j s_1 \tilde{\succ}_j s_\emptyset \ldots,$$

$$s_1 \tilde{\succ}_{i_1} s^* \tilde{\succ}_{i_1} s_\emptyset \ldots,$$

$$\mu_k \tilde{\succ}_k s_\emptyset \tilde{\succ}_k \ldots \quad \forall k \in I \backslash \{i, j, i_1\}.$$

Step 2: Fix any $s \in S$ and any $k \in \mu_s$. The construction of $\tilde{g}_{-i}$ guarantees $\tilde{g}_k^s = g_k^s$. Thus, $\pi(\mu, (g_i, \tilde{g}_{-i})) = \pi(\mu, g)$. Next, following a similar application procedure as in Case 3.1 (Step 2), we reach $DA_j(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = s_1$, $DA_{i_1}(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = s^*$ and $DA_k(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = \mu_k$ for all $k \in I \setminus \{j, i_1\}$. There is a unique cycle $\gamma = \{i_1 j, j i_1\}$ in $G^*(DA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})))$ and once this cycle is solved we obtain μ. In this instance,

[11] This is the only place, where we need a plausible scenario where a student does not consent. For a discussion see also Section 6).

all students except i receive their top choice in μ. The TP algorithm thus terminates and $EDA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}), \tilde{c}) = \mu$.

Step 3: The DA algorithm arrives at $DA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = s^*$, $DA_j(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = s_1$, $DA_{i_1}(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = s_\emptyset$ and $DA_k(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = \mu_k$ for all $k \in I \backslash \{i, j, i_1\}$. Notably, j is not eligible for s^*, since $\tilde{c}_{i_1} = 0$. Therefore, we cannot add ji to the graph and thus there is no cycle in $G^*(DA(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})))$. In conclusion, $EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = s^*$.

Case 3.2.2.2: $\pi_{\mu_i}(\mu, g) \neq 0$ and $g_i^{s^*} < \pi_{s^*}(\mu, g)$. That is, (C2) holds and we thus have $g_i^{s^*} < g_j^{s^*}$. Since $\pi_{s^*}(\mu, g) \neq 0$ and $\pi_{\mu_i}(\mu, g) \neq 0$, there are only three schools, namely s_1, μ_i, s^*, which exhaust their capacity under μ. In this last subcase, we show that statement (2) is satisfied.

First note that since i has justified envy for s_1 at μ, there exists a cycle containing i_1 that is solved in the TP process. Second, by non-wastefulness of EDA, if a school is contained in one solved cycle, it exhausts its capacity under the final matching. Recall that only s_1, μ_i, s^* exhaust their capacity at μ. Thus, the candidate student for forming a cycle can only be assigned to s^*. Therefore, we can construct exactly one cycle with i_1 and some $l \in \mu_{s^*}$.

Now select any $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$. Since $g_i^{s^*} < \pi_{s^*}(\mu, g)$ and by our arguments made above, we have $\tilde{g}_{i_1}^{s^*} > \tilde{g}_l^{s^*} > g_i^{s^*}$ and $DA_{i_1}(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = s^*$. However, this implies that i will be rejected by s^* under DA when she reports $\tilde{\succ}_i$. As a result, $DA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = \mu_i$ and statement (2) holds.

This completes the proof for the case in which $\tilde{S}$ is a singleton. To finish the proof, suppose now that $\tilde{S}$ contains multiple elements. We denote the top ranked school on $\tilde{\succ}_i$ among all schools in $\tilde{S}$ by s_1. Specifically, let $\succ_i^1$ be such that $s_1 \succ_i^1 \mu_i$ and $s \succ_i^1 s'$ if $s \succ_i s'$ for all $s, s' \in S \setminus \{s_1\}$. Since s_1 is the only permuted school on $\succ_i^1$ compared to $\succ_i$, we can apply the arguments above (for singleton $\tilde{S}$) to $\succ_i^1$. Here, we distinguish two cases. In the first case, suppose that the observation $(\mu, \pi(\mu, g))$ is such that statement (1) holds for $\succ_i^1$. That is, we find $(g_{-i}^1, \succ_{-i}^1, c_{-i}^1) \in \mathcal{I}(\mu, \succ_i^1, c_i)$

such that $EDA_i(g^1, (\succ_i^1, \succ_{-i}^1), c^1) = s_1$. Note that all our constructions above satisfy that $DA_i(g^1, (\succ_i^1, \succ_{-i}^1)) = EDA_i(g^1, (\succ_i^1, \succ_{-i}^1), c^1) = s_1$. Since $SU_{s_1}^{\tilde{\succ}_i} = SU_{s_1}^{\succ_i^1}$, we obtain $DA_i(g^1, (\tilde{\succ}_i, \succ_{-i}^1)) = EDA_i(g^1, (\tilde{\succ}_i, \succ_{-i}^1), c^1) = s_1$. Thus, we can conclude that statement (1) also holds for misreport $\tilde{\succ}_i$ for the first case. In the second case, suppose that the observation $(\mu, \pi(\mu, g))$ falls into the case where statement (2) holds for $\succ_i^1$. Then, we need further consider the second ranked school among $\tilde{S}$ on $\tilde{\succ}$, denoted by s_2. Specifically, we construct $\succ_i^2$ such that $s_1 \succ_i^2 s_2 \succ_i^2 \mu_i$ and $s \succ_i^2 s'$ if $s \succ_i s'$ for all $s, s' \in S \setminus \{s_1, s_2\}$. Since we assume that $\succ_i^1$ has no influence on the result at all, we can again apply the arguments for the singleton case to $\succ_i^2$. That is, we consider whether statement (1) or statement (2) applies to $\succ_i^2$. If statement (1) holds for $\succ_i^2$, then as explained above we can conclude that statement (1) holds for $\tilde{\succ}_i$. Otherwise, we further consider the third ranked school among $\tilde{S}$ on $\tilde{\succ}$. In the following, we iteratively apply the above arguments by adding a new school from $\tilde{S}$ through each iteration. Once we arrive at a step where statement (1) holds, we stop and conclude that statement (1) holds for $\tilde{\succ}_i$. On the contrary, if for all schools in $\tilde{S}$ the observation (2) holds, then we conclude that statement (2) holds for the misreport $\tilde{\succ}_i$. □

We move to the final class of misreports in which all schools that are truly less preferred to μ_i still rank lower than μ_i. That is, in the rest of the proof, we consider $\tilde{\succ}_i \in \mathcal{P}$ such that $SU_{\mu_i}^{\tilde{\succ}_i} \subseteq SU_{\mu_i}^{\succ_i}$ and for which there exists $s, s' \in SL_{\mu_i}^{\succ_i}$ such that $s \succ_i s'$ and $s' \tilde{\succ}_i s$. Our strategy is to show that if a student could have been improved upon truth through such a misreport $\tilde{\succ}_i$ in a plausible scenario, then the misreport could also have made the misreporting student worse off in another plausible scenario.

Before we formally show the above argument, we provide three auxiliary results. Throughout the remaining discussion, we fix some $(g'_{-i}, \succ'_{-i}, c'_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$. Henceforth, we use g' to refer to (g_i, g'_{-i}) and c' to refer to (c_i, c'_{-i}). Also, let for any $\succ'_i \in \mathcal{P}$ and any $s \in S$, the weak upper contour set of $\succ'_i$ at s be $U_s^{\succ'_i} = \{s' \in S \mid s' \succeq'_i s\}$.

Next, let $S' = \{s' \in SL_{\mu_i}^{\succ_i} | \; \exists \; \tilde{s} \in SL_{\mu_i}^{\succ_i} : s' \succ_i \tilde{s} \text{ and } \tilde{s} \; \tilde{\succ}_i \; s'\}$. Note that we now consider a misreport $\tilde{\succ}_i$ of the class where $SU_{\mu_i}^{\tilde{\succ}_i} \subseteq SU_{\mu_i}^{\succ_i}$ and hence according to Proposition 1, $EDA_i(g', (\tilde{\succ}_i, \succ'_{-i}), c') \neq EDA_i(g', (\succ_i, \succ'_{-i}), c')$ implies that S' must be non-empty. In the following, select any TP process with input $(g', (\succ_i, \succ'_{-i}), c')$ and denote it by $pTP^{\succ}$. Let $EDA^t(\succ)$ be the outcome of the t_{th} step in $pTP^{\succ}$. Then, we collect the set of schools to which i is (temporarily) assigned during $pTP^{\succ}$ in $S_i = \{\hat{s} \in S \mid \exists t \in \mathbb{N} : EDA_i^t(\succ) = \hat{s}\}$.

Lemma 6. *If $EDA_i(g', (\tilde{\succ}_i, \succ'_{-i}), c') \succ_i \mu_i$, then there exists $s' \in S'$ such that $g_i^{s'} > \pi_{s'}(\mu, g) > 0$.*

Proof. We prove the contrapositive statement. Note that in the process of TP algorithm, scores of assigned students are weakly decreasing at each school from step to step. Thus, for any $\hat{s} \in S_i$, we have $g_i^{\hat{s}} \geq \pi_{\hat{s}}(\mu, g)$. Also, schools in S_i must have positive cutoffs. Therefore, by assumption of S', we have $S' \cap S_i = \emptyset$. Hence, for any $\hat{s} \in S_i$, $SU_{\hat{s}}^{\tilde{\succ}_i} \subseteq SU_{\hat{s}}^{\succ_i}$. By Lemma 2, we reach $EDA_i(g', (\tilde{\succ}_i, \succ'_{-i}), c') = \mu_i$. This completes the proof. □

Lemma 7. *If $EDA_i(g', (\tilde{\succ}_i, \succ'_{-i}), c') \succ_i \mu_i$, then $\mu_i \succ_i DA_i(g', (\succ_i, \succ'_{-i}))$.*

Proof. Since EDA guarantees $\mu_i \succeq_i DA_i(g', (\succ_i, \succ'_{-i}))$, we assume by contradiction that $DA_i(g', (\succ_i, \succ'_{-i})) = \mu_i$. Recall that $\tilde{\succ}_i$ satisfies $SU_{\mu_i}^{\tilde{\succ}_i} \subseteq SU_{\mu_i}^{\succ_i}$. This assumption implies that for any $\hat{s} \in S_i$, $SU_{\hat{s}}^{\tilde{\succ}_i} \subseteq SU_{\hat{s}}^{\succ_i}$. By Lemma 2, we can infer $EDA_i(g', (\tilde{\succ}_i, \succ'_{-i}), c') = \mu_i$, which contradicts to $EDA_i(g', (\tilde{\succ}_i, \succ'_{-i}), c') \succ_i \mu_i$. □

Based on Lemma 7, we assume that $\mu_i \succ_i DA_i(g'_{-i}, (\succ_i, \succ'_{-i}))$ from now on. This implies that we have $\pi_{\mu_i}(\mu, g) \neq 0$. Moreover, by Lemma 6 there exists a maximal and non-empty set $S_1 \subseteq S'$ such that $s_1 \in S_1$ if and only if $g_i^{s_1} > \pi_{s_1}(\mu, g) > 0$. For the rest of the proof, let $r^* \in S_1$ be such that $r^* \succeq_i s_1$ for any $s_1 \in S_1$. Furthermore, we collect in $S_2 = \{s_2 \in L_{\mu_i}^{\succ_i} \mid r^* \succ_i s_2, s_2 \; \tilde{\succ}_i \; r^*\}$ and denote with $s^* \in S_2$ the school such that $s^* \; \tilde{\succeq}_i \; s_2$ for any $s_2 \in S_2$.

Lemma 8. *If $EDA_i(g', (\tilde{\succ}_i, \succ'_{-i}), c') \succ_i \mu_i$, then $\pi_{s^*}(\mu, g) \neq 0$.*

Proof. We show the contrapositive statement. That is, given $\pi_{s^*}(\mu, g) = 0$, we prove that $\mu_i \succeq_i EDA_i(g', (\tilde{\succ}_i, , \succ'_{-i}), c')$. Let $DA_i(g', (\succ_i, \succ'_{-i})) = \nu_i$. Since we assume $\pi_{\mu_i}(\mu, g) \neq 0$, it follows $\pi_{\nu_i}(\mu, g) \neq 0$. That is, $\nu_i \neq s^*$. In the following, we consider two cases that are distinguished by the relative ranking of s^* and ν_i on $\tilde{\succ}_i$.

In the first case, suppose $\nu_i \tilde{\succ}_i s^*$. Note that by the selection of s^* and the assumption $\nu_i \tilde{\succ}_i s^*$ we can infer that for any $\hat{s} \in S_i$, $SU_{\hat{s}}^{\tilde{\succ}_i} \subseteq SU_{\hat{s}}^{\succ_i}$ and thus $\mu_i = EDA_i(g', (\tilde{\succ}_i, , \succ'_{-i}), c')$ by Lemma 2.

In the second case, suppose $s^* \tilde{\succ}_i \nu_i$. We show $\mu_i \succ_i EDA_i(g', (\tilde{\succ}_i, \succ'_{-i}), c')$ here. We first argue $SU_{s^*}^{\tilde{\succ}_i} \subseteq SU_{\nu_i}^{\succ_i}$. By contradiction, suppose that there exists $r' \in S$ such that $r' \in SU_{s^*}^{\tilde{\succ}_i}$ and $r' \notin SU_{\nu_i}^{\succ_i}$. Then, we know (1) $\nu_i \succ_i r'$, (2) $r' \tilde{\succ}_i s^*$ and thus (3) $r' \tilde{\succ}_i \nu_i$. Since $g_i^{\nu_i} > \pi_{\nu_i}(\mu, g) > 0$, by (1) and (3) we can infer $\nu_i \in S_1$. Thus, the selection of r^* ensures that $r^* \succeq_i \nu_i$, which combined with (1) shows $r^* \succ_i r'$. Moreover, from (2) and the construction of S_2 we have $r' \tilde{\succ}_i s^* \tilde{\succ}_i r^*$. Note that $r^* \succeq_i \nu_i$ and $r' \tilde{\succ}_i s^* \tilde{\succ}_i r^*$ and we reach a contradiction to how s^* is selected. Thus, we have $SU_{s^*}^{\tilde{\succ}_i} \subseteq SU_{\nu_i}^{\succ_i}$. Next, since by assumption s^* has vacant seat at $EDA(g', (\succ_i, \succ'_{-i}), c')$, it also has vacant seat at $DA(g', (\succ_i, \succ'_{-i}))$. With the two findings above, we can use the arguments from Case 1 of Lemma 5 and conclude that no student strictly prefers $DA_i(g', (\tilde{\succ}_i, \succ'_{-i})) = s^*$ to her own assignments at $DA(g', (\tilde{\succ}_i, \succ'_{-i}))$. We then apply Lemma 4 and reach $EDA_i(g', (\tilde{\succ}_i, \succ'_{-i}), c') = s^*$. Since $\mu_i \succ_i s^*$, the proof is complete. □

We finally show that when i would have reported $\tilde{\succ}_i$, then she could have been worse off by being assigned to s^* in some plausible scenario.

Lemma 9. *If $EDA_i(g', (\tilde{\succ}_i, \succ'_{-i}), c') \succ_i \mu_i$, then there exists $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$ such that $\mu_i \succ_i EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = s^*$.*

Proof. Note that by Lemma 8, we only need to construct such a scenario for cases where $\pi_{s^*}(\mu, g) > 0$. Similar as in the proof of Lemma 5, we go through a series of standardized steps

Step 1: We construct a candidate scenario $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i})$.

Step 2: We show that $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$.

Step 3: We argue that $EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = s^*$.

Recall that $r^* \in S_1$ is the school that ranks highest on $\succ_i$ among all schools in S_1. Let $j \in I$ be an arbitrary student such that $\mu_j = s^*$, and let $l \in I$ be such that $\mu_l = r^*$ and $g_l^{r^*} = \pi_{r^*}(\mu, g)$. Moreover, consider the set $\bar{S} = \{s' \in SU_{r^*}^{\succ_i} \mid g_i^{s'} > \pi_{s'}(\mu, g)\}$ and denote $\bar{S} = \{s_1, s_2, \ldots, s_T\}$. Without loss of generality, let $s_1 \succ_i s_2 \succ_i \ldots \succ_i s_T$. Since $s^* \in SL_{r^*}^{\succ_i}$, we know that $s^* \notin \bar{S}$. For each $t \in \{1, ..., T\}$, denote the student with the lowest score assigned to s_t in μ by i_t and collect all such students in $\bar{I} = \{i_1, \ldots, i_T\}$. Since we already know that $\pi_{\mu_i}(\mu, g) \neq 0$ and $\pi_{s^*}(\mu, g) \neq 0$, it suffices to consider different cardinalities of $\bar{S}$ for distinguishing characteristic observations of student i.

Case 1: $|\bar{S}| \neq 1$**.** *Step 1:* We start with the candidate score structure. Let $\tilde{g}_{-i}$ be such that

- $\tilde{g}_l^{\mu_i} > \tilde{g}_j^{\mu_i} > g_i^{\mu_i}$; and $\tilde{g}_k^{\mu_i} = g_k^{\mu_i}$ for all $k \in I \setminus \{i, j, l\}$ and;
- $g_i^{r^*} > \tilde{g}_l^{r^*}$; and $\tilde{g}_k^{r^*} = g_k^{r^*}$ for all $k \in I \setminus \{i, j\}$ and;
- $\tilde{g}^{s'} = g^{s'}$ for any $s' \in S \setminus \{s_1, \ldots, s_T, \mu_i, r^*\}$.

Let $i_0 = i_T$ and $s_{T+1} = s_1$. In case that $\bar{S} \neq \emptyset$, for any $s_t \in \bar{S}$:

- $\tilde{g}_{i_{t-1}}^{s_t} > \tilde{g}_i^{s_t} > \tilde{g}_{i_t}^{s_t}$; and $\tilde{g}_k^{s_t} = g_k^{s_t}$ for all $k \in I \setminus \{i, i_{t-1}\}$.

Next, we specify $\tilde{c}_{-i}$ such that for all $i' \in I \setminus \{i\}$ it holds that $\tilde{c}_{i'} = 1$ and consider preference profile $\tilde{\succ}_{-i} \in \mathcal{P}_{-i}$:

$$s_t \mathrel{\tilde{\succ}_{i_t}} s_{t+1} \mathrel{\tilde{\succ}_{i_t}} s_\emptyset \mathrel{\tilde{\succ}_{i_t}} \dots \quad \forall t \in \{1, \dots, T\},$$

$$r^* \mathrel{\tilde{\succ}_l} \mu_i \mathrel{\tilde{\succ}_l} s_\emptyset \mathrel{\tilde{\succ}_l} \dots\ ,$$

$$\mu_i \mathrel{\tilde{\succ}_j} s^* \mathrel{\tilde{\succ}_j} s_\emptyset \mathrel{\tilde{\succ}_j} \dots\ ,$$

$$\mu_k \mathrel{\tilde{\succ}_k} s_\emptyset \mathrel{\tilde{\succ}_k} \dots \quad \forall k \in I \backslash (\bar{I} \cup \{i, j, l\}).$$

Step 2: The construction of $\tilde{g}_{-i}$ ensures that $\pi(\mu, (g_i, \tilde{g}_{-i})) = \pi(\mu, g)$. For the constructed scenario DA leads to $DA_i(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = r^*$, $DA_j(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = s^*$, $DA_l(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = \mu_i$, $DA_{i_t}(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = s_{t+1}$ for each $t \in \{1, \dots, T\}$ and $DA_k(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})) = \mu_k$ for $k \in I \backslash (\bar{I} \cup \{i, j, l\})$. Consider the corresponding application process. At the first step, for all $k \in I \setminus (\bar{I} \cup \{i, j, l\})$, k is accepted at μ_k, j is accepted at μ_i, l is accepted at r^*, and for all $t \in \{1, \dots, T\}$, i_t is accepted at s_t. If i's top choice is not s_1, let $t_1 \in \mathbb{N}$ be the step, in which i applies to s_1 and is tentatively accepted. In all the previous steps $t < t_1$, student i is rejected. For each $t \in \{1, \dots, T\}$, this leads to i_t getting rejected by s_t and applying to s_{t+1} in the next step, causing i_{t+1} being rejected by s_{t+1} and so forth. Eventually i is rejected by s_1 at step $t_1 + T$. Then, student i is rejected at the remaining schools in $SU_{r^*}^{\succ_i}$ until being accepted at r^*, in favor of l. Student l then applies to μ_i such that j gets rejected. Next, j applies to s^* and gets accepted. Here, the algorithm terminates.

We now show that the cycle selection under a TP process ends in the observed matching μ. Since j is permanently matched in $DA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}))$ and $\tilde{c}_j = 1$, we know that $G^*(DA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i})))$ contains cycle $\gamma^1 = \{il, li\}$ and solving it yields $EDA^1(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}), \tilde{c}) = \gamma^1 \circ DA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}))$, where compared to $DA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}))$, only i and l switch their assignments.

Next, since $c_i = 1$ and i is permanently matched to μ_i in $EDA^1(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}), \tilde{c})$, whenever $\bar{S}$ is non-empty, $G^*(EDA^1(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}), \tilde{c}))$ contains a unique cycle

$$\gamma^2 = \{i_T i_{T-1}, i_{T-1} i_{T-2}, ..., i_{t+1} i_t, ... i_2 i_1, i_1 i_T\}$$

which once solved yields matching μ. Since all students except i and j get their top-choice, and both i, j are permanently matched, there is no cycle in $G^*(\mu)$. Therefore, $EDA(\tilde{g}, (\succ_i, \tilde{\succ}_{-i}), \tilde{c}) = \mu$.

Step 3: Reviewing the application process above, we get $DA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = s^*$. Moreover, note that apart from the students who are matched with school s^* at $DA(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}))$, student j is the only one who ranks s^* above $s_\emptyset$ in $\tilde{\succ}_{-i}$. However, notice that $DA_j(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i})) = \mu_i \tilde{\succ}_j s^*$ and school s^* is underdemanded in $DA(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}))$. By Lemma 4, we can infer $EDA_i(\tilde{g}, (\tilde{\succ}_i, \tilde{\succ}_{-i}), \tilde{c}) = s^*$. This completes the proof for Case 1.

Case 2: $|\bar{S}| = 1$**.** *Step 1:* Let $\tilde{g}_{-i}$ be such that

- $\tilde{g}_l^{s_1} > g_i^{s_1} > \tilde{g}_{i_1}^{s_1}$; and $\tilde{g}_k^{s_1} = g_k^{s_1}$ for all $k \in I \setminus \{i, l\}$ and;
- $\tilde{g}_{i_1}^{\mu_i} > \tilde{g}_j^{\mu_i} > g_i^{\mu_i}$; and $\tilde{g}_k^{\mu_i} = g_k^{\mu_i}$ for all $k \in I \setminus \{i, j, i_1\}$ and;
- $g_i^{r^*} > \tilde{g}_{i_1}^{r^*} > \tilde{g}_l^{r^*}$; and $\tilde{g}_k^{r^*} = g_k^{r^*}$ for all $k \in I \setminus \{i, i_1\}$ and;
- $\tilde{g}^{s'} = g^{s'}$ for any $s' \in S \setminus \{s_1, \mu_i, r^*\}$.

Under $\tilde{c}_{-i}$, let for all $i' \in I \setminus \{i\}$ be $\tilde{c}_{i'} = 1$ and let $\tilde{\succ}_{-i} \in \mathcal{P}_{-i}$ be:

$$s_1 \tilde{\succ}_{i_1} r^* \tilde{\succ}_{i_1} \mu_i \tilde{\succ}_{i_1} s_\emptyset \tilde{\succ}_{i_1} \ \ldots,$$

$$r^* \tilde{\succ}_l s_1 \tilde{\succ}_l s_\emptyset \tilde{\succ}_l \ \ldots \ ,$$

$$\mu_i \tilde{\succ}_j s^* \tilde{\succ}_j s_\emptyset \tilde{\succ}_j \ \ldots \ ,$$

$$\mu_k \tilde{\succ}_k s_\emptyset \tilde{\succ}_k \ \ldots \ \ \forall k \in I \setminus \{i, j, l, i_1\}.$$

Step 2 and Step 3: Here, we can almost resemble the arguments in Step 2 and Step 3 for Case 1. That is, i is worse off by being finally assigned to s^*, which is underdemanded under the DA outcome. $\square$

Since the conclusion holds for any observation, any student and any problem, we conclude that EDA is regret-free truth-telling.

Appendix D Proof of Proposition 2

With a similar technique as in the proof of Proposition 1 in Fernandez (2020). Fix an arbitrary problem $(I, S, q, g, \succ, c)$ and fix an arbitrary $i \in I$. We divide the set of possible misreports into three exhaustive cases.

Case 1 Let under $\succ_i'$ exists $s \in S$ such that $s_\emptyset \succ_i s$ and $s \succ_i' s_\emptyset$. Let i submit $\succ_i'$ and consider the pair $(\mu, \pi(\mu, g))$ such that $\mu_i = s$ and $g_i^{s'} < \pi_{s'}(\mu, g)$ for all $s' \in SU_s^{\succ_i'}$. At first, we show that $\mu \in \mathcal{M}|_{(\succ_i', c_i)}$ by constructing $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i})$ that leads to $(\mu, \pi(\mu, g))$: That is, we show that $(\mu, \pi(\mu, g))$ is an observation under EDA. Let $\tilde{g}_{-i}$ be such that, for each $s' \in SU_{\mu_i}^{\succ_i'}$, each student in $\mu_{s'}$ is among the top $q_{s'}$'s scored students at school s'. Let i rank highest on $\tilde{g}^s$ and let the remaining scores be arbitrary. Let $\tilde{\succ}_{-i}$ be such that for each $j \in I \setminus \{i\}$, $\tilde{\succ}_j$ only ranks μ_j as acceptable and assume $\tilde{c} = c$. Apparently, we have $\pi(\mu, (g_i, \tilde{g}_{-i})) = \pi(\mu, g)$ and $EDA(\tilde{g}, (\succ_i', \tilde{\succ}_{-i}), \tilde{c}) = \mu$. Thus, $\mu \in \mathcal{M}|_{(\succ_i', c_i)}$. Now note that for any $(\hat{g}_{-i}, \hat{\succ}_{-i}, \hat{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$ it holds that $EDA_i(\hat{g}, (\succ_i, \hat{\succ}_{-i}), \hat{c}) \succeq_i s_\emptyset$, since EDA is individually rational. Since $s_\emptyset \succ_i s$, student i regrets $\succ_i'$ through $\succ_i$ at $(\mu, \pi(\mu, g))$.

Case 2 Let for $\succ_i'$ exist $s \in S$ such that $s_\emptyset \succ_i' s$ and $s \succ_i s_\emptyset$. Suppose i submits $\succ_i'$ and consider $(\mu, \pi(\mu, g))$ such that $\mu_i = s_\emptyset$, $\pi_s(\mu, g) = 0$ and $g_i^{s'} < \pi_{s'}(\mu, g)$ for all $s' \in SU_{s_\emptyset}^{\succ_i'}$. Notably, by doing the same construction $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i})$ as in Case 1, we can infer $\mu \in \mathcal{M}|_{(\succ_i', c_i)}$. Next, note that EDA is non-wasteful and as such for any

$(\hat{g}_{-i}, \hat{\succ}_{-i}, \hat{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$, it holds that $EDA_i(\hat{g}, (\succ_i, \hat{\succ}_{-i}), \hat{c}) = s$. Since $s \succ_i s_\emptyset$, student i regrets $\succ'_i$ through $\succ_i$ at $(\mu, \pi(\mu, g))$.

Case 3 Consider $\succ'_i$ which only contains variations in the acceptable and unacceptable set. For any $\succ''_i \in \mathcal{P}$, collect in $A_i(\succ''_i)$ all acceptable schools. The following labeling for any $\succ''_i \in \mathcal{P}$ in the acceptable set $A_i(\succ''_i)$ ensures that a school's index corresponds to its position in $\succ''_i$. Precisely, we denote s''_1 as the $\succ''_i$-maximal element on $A_{i,1}(\succ''_i) = A_i(\succ''_i)$ and s''_2 as the $\succ''_i$-maximal element on $A_{i,2}(\succ''_i) = A_{i,1}(\succ''_i) \setminus \{s''_1\}$, and so forth. Let $|A_i(\succ_i)| = N \in \mathbb{N}$ be the number of acceptable schools under $\succ_i$ and consider $\succ'_i$ as described above. Since $\succ'_i$ is a variation, there exists $n^* = \arg\min_n \{n \leq N \mid s'_n \neq s_n\}$. Next, let student i observe $(\mu, \pi(\mu, g))$ such that $\mu_i = s'_{n^*}$, $\pi_{s_{n^*}}(\mu, g) = 0$ and $g_i^{s'} < \pi_{s'}(\mu, g)$ for all $s' \in SU_{s'_{n^*}}^{\succ'_i}$. Again, by doing the same construction $(\tilde{g}_{-i}, \tilde{\succ}_{-i}, \tilde{c}_{-i})$ as in Case 1, we can infer $\mu \in \mathcal{M}|_{(\succ'_i, c_i)}$.

Next, since s_{n^*} has capacity left, if i had reported $\succ_i$ then, for any $(\hat{g}_{-i}, \hat{\succ}_{-i}, \hat{c}_{-i}) \in \mathcal{I}(\mu, \succ_i, c_i)$, i would had been matched to s_{n^*}. Since $s_{n^*} \succ_i s'_{n^*}$, we conclude that i regrets $\succ'_i$ through $\succ_i$ at $(\mu, \pi(\mu, g))$. This completes the proof. $\square$

Appendix E Proposition 3

We aim at constructing a regret-free truth-telling stable dominating rule f that is neither stable nor efficient. Concretely, let f select the DA outcome except for a problem $(I, S, q, \hat{g}, \hat{\succ})$ as it is described in the proof of Theorem 2. In this problem, we have $S = \{s_1, s_2\}$, where both schools have unit capacity and $I = \{i_1, i_2, i_3\}$. Student i_1's preferences are $s_2 \; \hat{\succ}_{i_1} \; s_\emptyset \; \hat{\succ}_{i_1} \; s_1$, student i_2's preferences are $s_1 \; \hat{\succ}_{i_2} \; s_2 \; \hat{\succ}_{i_2} \; s_\emptyset$ and student i_3's preferences are $s_2 \; \hat{\succ}_{i_3} \; s_1 \; \hat{\succ}_{i_3} \; s_\emptyset$ and the score structure $\hat{g}$ satisfies $\hat{g}_{i_1}^{s_1} > \hat{g}_{i_3}^{s_1} > \hat{g}_{i_2}^{s_1}$ and $\hat{g}_{i_2}^{s_2} > \hat{g}_{i_1}^{s_2} > \hat{g}_{i_3}^{s_2}$. Let f select the efficient and non-stable matching $\hat{\mu} = \{(i_1, s_\emptyset), (i_2, s_1), (i_3, s_2)\}$ in this problem.[12] Since f always selects the DA outcome

[12]Notably, our argument extends directly to any rule f' that selects the DA outcome except for problems $(I, S, q, g', \hat{\succ})$ where g' share the same rankings as $\hat{g}$ (with different scores). For ease of

in problems with primitives other than (I, S, q) and since DA is regret-free truth-telling, it suffices to show that f is regret-free truth-telling in problems with (I, S, q). In what follows, we thus consider only problems with primitives (I, S, q).

We first consider i_1. Under f, for any pair of scores and preferences $(g, \succ) \in \mathcal{G}_I \times \mathcal{P}_I$, i_1 receives her most preferred school among schools she can be matched to in a stable matching. Note that this includes input $(\hat{g}, \hat{\succ})$, where i_1 receives $s_\emptyset = DA_{i_1}(\hat{g}, \hat{\succ})$. Thus, i_1 cannot improve by misreporting and hence does not regret telling the truth.

We next consider i_2 and i_3. Since i_2 and i_3 receive their top choices under $f(\hat{g}, \hat{\succ})$, both of them do not regret telling the truth for input $(\hat{g}, \hat{\succ})$. In the following, consider an arbitrary input $(g = \{g_i\}_{i \in I}, \succ = \{\succ_i\}_{i \in I})$ and let $(\mu, \pi(\mu, g))$ be the observation under $f(g, \succ)$. We first show that i_2 will not regret reporting her true preference $\succ_{i_2}$ under $(\mu, \pi(\mu, g))$. Concretely, suppose that i_2 improves upon μ_{i_2} under f by misreporting. By strategy-proofness of DA and the fact that f selects an outcome different from DA only if the input is $(\hat{g}, \hat{\succ})$, it follows that (1) i_2 misreports $\tilde{\succ}_{i_2} = \hat{\succ}_{i_2} \neq \succ_{i_2}$, (2) $g_{i_2} = \hat{g}_{i_2}$, (3) $(\hat{g}_{-i_2}, \hat{\succ}_{-i_2}) \in \mathcal{I}(\mu, \succ_{i_2})$. Observe that (2) and (3) imply that $\succ_{i_2}$ cannot have s_2 as the top choice, since i_2 would have been assigned to s_2 under μ. However, then, i_2 could never improve upon μ_{i_2} from misreporting. The same argument holds for $s_\emptyset$. Hence together with (1) we reach $s_1 \succ_{i_2} s_\emptyset \succ_{i_2} s_2$ and since i_2 cannot be matched to her top choice under μ, we have $\mu_{i_2} \neq s_1$. With $\succ_{i_2}$ and given (2) and (3), we know that $\mu_{i_1} = s_2$, $\mu_{i_2} = s_\emptyset$ and $\mu_{i_3} = s_1$. Accordingly, we have $\pi_{s_1}(\mu, g) = \hat{g}_{i_3}^{s_1} = g_{i_3}^{s_1}$ and $\pi_{s_2}(\mu, g) = \hat{g}_{i_1}^{s_2} = g_{i_1}^{s_2}$. Now, consider $\succ_{i_3}^*$: $s_1 \succ_{i_3}^* s_\emptyset \succ_{i_3}^* s_2$. Note that $(\hat{g}_{-i_2}, (\hat{\succ}_{i_1}, \succ_{i_3}^*)) \in \mathcal{I}(\mu, \succ_{i_2})$ and that $f_{i_2}((g_{i_2}, \hat{g}_{-i_2}), (\hat{\succ}_{i_1}, \tilde{\succ}_{i_2}, \succ_{i_3}^*)) = s_2$ and since $s_\emptyset \succ_{i_2} s_2$, student i_2 does not regret truthful-telling under $(\mu, \pi(\mu, g))$.

Next, suppose that i_3 improves by misreporting. We use a similar argument as for i_2 to reach that i_3's improvement upon μ_{i_3} would require $s_2 \succ_{i_3} s_\emptyset \succ_{i_3} s_1$: By strategy-proofness of DA and since f selects an outcome different from DA only if

presentation, we consider in the proof f that only selects a non-stable outcome for this specific problem $(I, S, q, \hat{g}, \hat{\succ})$.

the input is $(\hat{g}, \hat{\succ})$, i_3's improvement needs that (1′) i_3 misreports $\tilde{\succ}_{i_3} = \hat{\succ}_{i_3} \neq \succ_{i_3}$, (2′) $g_{i_3} = \hat{g}_{i_3}$, (3′) $(\hat{g}_{-i_3}, \hat{\succ}_{-i_3}) \in \mathcal{I}(\mu, \succ_{i_2})$. Conditions (2′) and (3′) imply that s_1 and $s_\emptyset$ cannot be top choices on $\succ_{i_3}$ and since i_3 must be able to improve, we also have $\mu_{i_3} \neq s_2$. Next, given $\succ_{i_3}$ under (2′) and (3′), we reach $\mu_{i_1} = s_2$, $\mu_{i_2} = s_1$ and $\mu_{i_3} = s_\emptyset$ while $\pi_{s_1}(\mu, g) = \hat{g}^{s_1}_{i_2} = g^{s_1}_{i_2}$ and $\pi_{s_2}(\mu, g) = \hat{g}^{s_2}_{i_1} = g^{s_2}_{i_1}$. However, consider $\succ^*_{i_2}$, where $s_1 \succ^*_{i_2} s_\emptyset \succ^*_{i_2} s_2$. Note that $(\hat{g}_{-i_3}, (\hat{\succ}_{i_1}, \succ^*_{i_2})) \in \mathcal{I}(\mu, \succ_{i_3})$ and $f_{i_3}((g_{i_3}, \hat{g}_{-i_3}), (\hat{\succ}_{i_1}, \succ^*_{i_2}, \tilde{\succ}_{i_3})) = s_1$. Since $s_\emptyset \succ_{i_3} s_1$, we reach that i_3 does not regret truth-telling under $(\mu, \pi(\mu, g))$.

Since there is no student who regrets being truthful, this completes the proof.

References

Atila Abdulkadiroğlu and Tayfun Sönmez. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747, 2003.

Atila Abdulkadiroğlu, Parag A. Pathak, and Alvin E. Roth. Strategy-proofness versus efficiency in matching with indifferences: Redesigning the nyc high school match. *American Economic Review*, 99(5):1954–78, 2009.

Samson Alva and Vikram Manjunath. Stable-dominating rules. Working paper, University of Ottawa, 2019a.

Samson Alva and Vikram Manjunath. Strategy-proof pareto-improvement. *Journal of Economic Theory*, 181:121–142, 2019b.

Eduardo M Azevedo and Jacob D Leshno. A supply and demand framework for two-sided matching markets. *Journal of Political Economy*, 124(5):1235–1268, 2016.

Michel Balinski and Tayfun Sönmez. A tale of two mechanisms: student placement. *Journal of Economic theory*, 84(1):73–94, 1999.

Keisuke Bando. On the existence of a strictly strong nash equilibrium under the student-optimal deferred acceptance algorithm. *Games and Economic Behavior*, 87:269 – 287, 2014.

Claudia Cerrone, Yoan Hermstrüwer, and Onur Kesten. School choice with consent: an experiment. *MPI Collective Goods Discussion Paper*, (2022/2), 2022.

Peter Chen, Michael Egesdal, Marek Pycia, and M Bumin Yenmez. Quantile stable mechanisms, 2014.

Benoit Decerf and Martin Van der Linden. Manipulability in school choice. *Journal of Economic Theory*, 197:105313, 2021.

Bnaya Dreyfuss, Ori Heffetz, and Matthew Rabin. Expectations-based loss aversion may help explain seemingly dominated choices in strategy-proof mechanisms. Technical report, National Bureau of Economic Research, 2019.

Lester E Dubins and David A Freedman. Machiavelli and the gale-shapley algorithm. *The American Mathematical Monthly*, 88(7):485–494, 1981.

Umut Dur, A Arda Gitmez, and Özgür Yılmaz. School choice under partial fairness. *Theoretical Economics*, 14(4):1309–1346, 2019.

Lars Ehlers. Truncation strategies in matching markets. *Mathematics of Operations Research*, 33(2):327–335, 2008.

Lars Ehlers and Thayer Morrill. (il) legal assignments in school choice. *The Review of Economic Studies*, 87(4):1837–1875, 2020.

Richard Engelbrecht-Wiggans. The effect of regret on optimal bidding in auctions. *Management Science*, 35(6):685–692, 1989.

Aytek Erdil and Haluk Ergin. What's the matter with tie-breaking? improving efficiency in school choice. *American Economic Review*, 98(3):669–89, 2008.

MA Fernandez. Deferred acceptance and regret-free truthtelling. Working paper, John Hopkins University, 2020.

Emel Filiz-Ozbay and Erkut Y Ozbay. Auctions with anticipated regret: Theory and experiment. *American Economic Review*, 97(4):1407–1418, 2007.

David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

Thomas Gilovich and Victoria Husted Medvec. The experience of regret: what, when, and why. *Psychological review*, 102(2):379, 1995.

Onur Kesten. School choice with consent. *The Quarterly Journal of Economics*, 125 (3):1297–1348, 2010.

Bettina Klaus and Flip Klijn. Median stable matching for college admissions. *International Journal of Game Theory*, 34(1):1, 2006.

Fuhito Kojima and Mihai Manea. Axioms for deferred acceptance. *Econometrica*, 78 (2):633–653, 2010.

Hyukjun Kwon and Ran I. Shorrer. Justified-envy-minimal efficient mechanisms for priority-based matching. *Microeconomics: Welfare Economics Collective Decision-Making eJournal*, 2019.

Vincent Meisner and Jonas von Wangenheim. School choice and loss aversion. *Available at SSRN 3985777*, 2021.

Siqi Pan. The instability of matching with overconfident agents. *Games and Economic Behavior*, 113:396–415, 2019.

Philip J Reny. Efficient matching in the school choice problem. Technical report, Working Paper, University of Chicago, 2021.

Alvin E Roth. Incentive compatibility in a market with indivisible goods. *Economics letters*, 9(2):127–132, 1982.

Alvin E Roth and Uriel G Rothblum. Truncation strategies in matching markets—in search of advice for participants. *Econometrica*, 67(1):21–43, 1999.

Lloyd Shapley and Herbert Scarf. On cores and indivisibility. *Journal of Mathematical Economics*, 1(1):23–37, 1974.

Qianfeng Tang and Jingsheng Yu. A new perspective on kesten's school choice with consent idea. *Journal of Economic Theory*, 154:543–561, 2014.

Chung-Piaw Teo and Jay Sethuraman. The geometry of fractional stable matchings and its applications. *Mathematics of Operations Research*, 23(4):874–891, 1998.

Peter Troyan and Thayer Morrill. Obvious manipulations. *Journal of Economic Theory*, 185:104970, 2020.

Peter Troyan, David Delacrétaz, and Andrew Kloosterman. Essentially stable matchings. *Games and Economic Behavior*, 120:370–390, 2020.

Robert Wilson. Game-theoretic analyses of trading processes. In *Advances in Economic Theory, Fifth World Congress*, pages 33–70, 1987.

Marcel Zeelenberg and Rik Pieters. A theory of regret regulation 1.0. *Journal of Consumer psychology*, 17(1):3–18, 2007.