UTMD Working Paper

The University of Tokyo
Market Design Center

UTMD-048

# Ekkyo Matching: How to Connect Separate Matching Markets for Welfare Improvement

Yuichiro Kamada
University of California, Berkeley

Fuhito Kojima
University of Tokyo

September 28, 2022

# EKKYO MATCHING: HOW TO CONNECT SEPARATE MATCHING MARKETS FOR WELFARE IMPROVEMENT

YUICHIRO KAMADA AND FUHITO KOJIMA

Abstract. We consider a school-choice matching model that allows for inter-district transfer of students, with the "balancedness" constraint: each student and school belongs to a region, and a matching is said to be balanced if, for each region, the outflow of students from that region to other regions is equal to the inflow of students from the latter to the former. Using a directed bipartite graph defined on students and schools, we characterize the set of Pareto efficient matchings among those that are individually rational, balanced and fair. We also provide a polynomial-time algorithm to compute such matchings. The outcome of this algorithm weakly improves student welfare upon the one induced when each region independently organizes a standard matching mechanism.

*JEL Classification Numbers*: C70, D47, D61, D63.

*Keywords*: ekkyo, balancedness, fairness, algorithm, efficiency, daycare allocation, school choice, matching with constraints

---

*Date*: September 28, 2022.

Kamada: Haas School of Business, University of California, Berkeley, Berkeley, CA 94720, y.cam.24@gmail.com. Kojima: Department of Economics, University of Tokyo, Tokyo, Japan, fuhitokojima1979@gmail.com. We thank Bobak Pakzad-Hurson and Shunya Noda for helpful comments. Yutaro Akita, Nanami Aoi, Yusuke Iwase, Kyohei Okumura, Akira Matsushita, and Ayano Yago provided excellent research assistance. Fuhito Kojima is supported by the JSPS KAKENHI Grant-In-Aid 21H04979.

Contents

## 1. Introduction

Matching theory has been applied in numerous real-life markets with a purpose of centralizing market transactions, but the centralized clearinghouses are still often organized at a (often small) local level. As a consequence, some efficiency gains are being missed. In Japan, for instance, allocation of slots at accredited daycares are conducted by individual municipal governments and, with few exceptions, a child can only attend a daycare in the municipality of their residence. The City of Tokyo, for example, is divided into 23 small municipalities with the average size being about a half of Manhattan's, and each conducts a matching independently.[1] Due to the small sizes of the regions, many families would find inter-district admissions—which is called the *ekkyo* admission and it is not prohibited by the law—to be a viable option. Moreover, as a large metropolitan area, many people cross a city boundary to commute, making it potentially more convenient to put their children to a daycare center close to their workplace but prohibited.[2]

We study how to improve upon mechanisms organized at the local level and achieve outcomes that have desirable fairness and efficiency properties. To do so, we depart from the standard model of matching between students and schools (Abdulkadiroğlu and Sönmez, 2003) by assuming that each school belongs to exactly one region while each student is a resident of exactly one region. We consider a balancedness constraint that requires that, for each region, the number of residents of other regions who are matched to a school in it, called the inflow, must be equal to the number of its residents who are matched to a school in other regions, called the outflow. Why is this a reasonable requirement? In the context of Japanese daycare allocation, for instance, each municipality heavily subsidizes daycares, so enrolling residents from other municipalities can be a severe financial burden. Our balancedness constraint is meant to alleviate municipalities' concerns by guaranteeing that no municipality carries an excessive burden.

We find that there does not always exist a balanced matching that satisfies stability, a standard desideratum in school choice literature that is equivalent to individual rationality, fairness, and non-wastefulness. Given this observation, we weaken our desideratum to only require individual rationality, balancedness and fairness. Among the matchings satisfying all the three conditions (there always exists such a matching[3]), which we call the *iBF*s,

[1]The average area of each of the 23 municipalities is 10.4 square miles, while Manhattan's area is 22.7 square miles.

[2]To get a sense of the magnitude, over 3.5 million people get on or get off a train in Tokyo's Shinjuku Station *each day*.

[3]All the conditions are satisfied by the matching where no student is matched to any school.

we focus on the ones that cannot be further improved upon in terms of students' welfare, which we call *efficient iBFs*.

We first characterize efficient iBFs. To this end, we define a novel bipartite graph called a fair improvement graph (FIG henceforth) on a matching, where the vertices on one side represent the students and those on the other side represent the schools, and existence of an arrow between two vertices depends on preferences, priorities, and the given matching. We show that an iBF is an efficient iBF if and only if there exists no "FIG cycle," a cycle on the FIG, for the matching.

Based on our characterization of efficient iBF, we then provide a polynomial-time algorithm that finds an efficient iBF. The algorithm is called the FIG cycle algorithm and is illustrated in some detail with an example in Section 1.1. Roughly, each step of the FIG algorithm checks if the current iBF allows for a FIG cycle and, if so, finds a relocation of students that improves outcomes for students while retaining individual rationality, balancedness and fairness.

To understand the markets with the balancedness constraint, we provide further discussions. First, we examine the strategy properties. We show that our mechanism based on the FIG cycle algorithm turns out not to be strategy-proof. We, however, also show that there is no strategy-proof mechanism that always outputs an efficient iBF. Second, comparative statics are provided to evaluate the effect of merging and splitting regions. Third, we consider the case with weak priority, which often arises in school choice applications. Finally, we review the related literature.

We would like to emphasize that our analysis is applicable beyond daycare allocation. Also in Japan, choice systems for elementary and secondary schools are organized at the small municipal level as well. Naturally, there exist much potential demand for enrolling in schools in other municipalities, but admission is severely limited. The issue is not limited to Japan either. In the U.S., for example, school choice is basically organized at a highly local level, but some form of interdistrict school choice is practiced in 43 States (Education Commission of the States, 2017). The analysis of our paper, and particularly the FIG cycles algorithm, could be applied to improve efficiency of school choice mechanisms.

At a high level, the highly local nature of resource allocation is widespread beyond daycare allocation or school choice. In kidney exchange in the U.S., for instance, individual transplant centers often conduct exchanges on their own before sending remaining participants to national exchange, resulting in significant efficiency loss (Agarwal et al., 2019). In COVID-19 vaccine allocation in Japan in 2021, individual municipalities were charged with vaccinating their respective residents, which resulted in situations in which

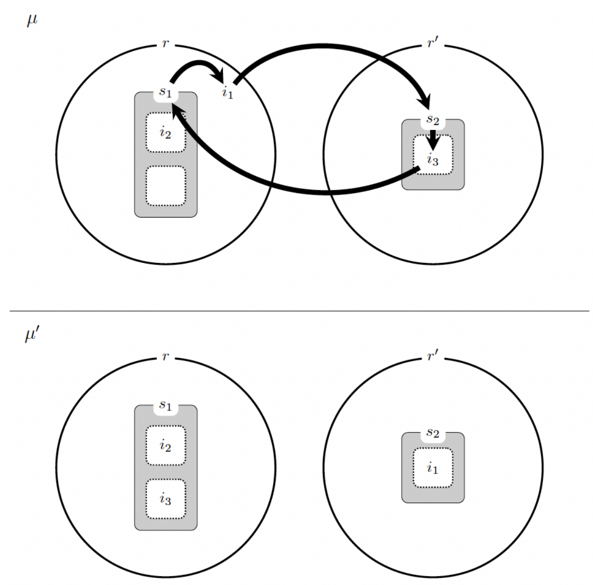


FIGURE 1

vaccine stocks run out quickly in one municipality while extra stocks remain unused in another. We envision that research is called for to understand how to overcome inefficiency from the localized nature of allocation problem in a practical manner when existing legal, institutional and other constraints prohibit full integration.

1.1. **Illustrative Example.** In this paper we introduce an algorithm of inter-region transfer that improves students' welfare while respecting the balancedness condition and fairness. To gain intuition for why an inter-region transfer may improve welfare, consider the following simple environment (see Figure 1 for an illustration). There are two regions, $r$ and $r'$. One school $s_1$ and two students $i_1$ and $i_2$ reside in $r$ while one school $s_2$ and one student $i_3$ reside in $r'$. School $s_1$ has the capacity of two while school $s_2$ has the capacity of one. Student preferences and school preferences are given as follows:

$$\succ_{i_1}: s_2, \qquad\qquad \succ_{s_1}: i_1, i_2, i_3$$

$$\succ_{i_2}: s_2, s_1, \qquad\qquad \succ_{s_2}: i_3, i_1, i_2.$$

$$\succ_{i_3}: s_1, s_2,$$

If an assignment of students to schools is determined region by region and there is no inter-region transfer, the efficient matching is given by

$$\mu = \begin{pmatrix} s_1 & s_2 & \emptyset \\ i_2 & i_3, & i_1 \end{pmatrix},$$

which is realized by, for instance, running the student-proposing deferred acceptance algorithm in each region separately. However, if the two students $i_1$ and $i_3$ are sent to the schools in each other's regions, that is, if $i_1$ goes to $s_2$ and $i_3$ goes to $s_1$, then in the realized matching,

$$\mu' = \begin{pmatrix} s_1 & s_2 & \emptyset \\ i_2, i_3 & i_1 & \emptyset \end{pmatrix},$$

students are matched with their respective first-best school. In doing so, region $r_1$ takes in one student from outside (student $i_3$) while sending one student outside (student $i_1$), so the balancedness condition is satisfied.

Two things are noteworthy here. First, the number of students who are matched to some school is increased from $\mu$ to $\mu'$. This is because, by swapping students between the two regions, the unmatched student $i_1$ was able to be matched with a school. We introduce an algorithm to make such an improvement possible.[4]

Second, there is another matching that respects the balancedness condition, which is

$$\mu'' = \begin{pmatrix} s_1 & s_2 & \emptyset \\ i_3 & i_2 & i_1 \end{pmatrix}.$$

However, this matching is not fair according to our definition, as $i_1$ is ranked higher than $i_3$ at $s_2$. This suggests that care is needed about who can be moved to new schools across regions. The algorithm we introduce ensures that fairness is respected when an improvement is made.

We aim to achieve efficiency via inter-regional transfer like the one described in the above example. Specifically, we propose an algorithm that takes as an input an arbitrary iBF and achieves a Pareto improvement. The algorithm is based on a directed bipartite graph between students and schools that we call the *fair improvement graph* (or the *FIG*), and in each of its steps it "implements" a cycle in this graph—called a *FIG cycle*—, i.e., we move a student to a school that she points to. The outcome of repeatedly implementing FIG cycles turns out to be an *efficient iBF*. In fact, our main results characterize efficient iBF using cycles: we show that *an iBF is an efficient iBF if and only if there is no FIG cycle on it.*

Let us now discuss these results in the context of the aforementioned example. In our FIG, each school is pointed to by the top student (according to its priority) among those who strictly prefer the school to their current match and are acceptable to the school. This means that, under $\mu$, $s_1$ is pointed to by $i_3$ while $s_2$ is pointed to by $i_1$. Note that

[4] As we will explain in more detail later, this type of improvement is in a sharp contrast with most existing literature where the number of matched students is constant between before and after an improvement.

$i_2$ cannot point to $s_2$ as he is not the "top student" for $s_2$ ($i_1$ is). This pointing rule is the same as that of Erdil and Ergin (2008). Also, each school points to the students that are matched to the school. So, $s_1$ points to $i_2$ and $s_2$ points to $i_3$. This pointing rule is common in the algorithms that use cycles such as Top Trading Cycles (TTC) or Erdil and Ergin (2008).

Note, however, that the graph constructed in this way does not have a cycle even though matching $\mu$ is not an efficient iBF In order to achieve a Pareto improvement, we additionally require that each school with a vacancy points to any student matched to another school in the school's region as well as any unmatched students living in that region. In our example, this lets $s_1$ point to $i_1$.[5] With this, there is a cycle "$i_1 \to s_2 \to i_3 \to s_1 \to i_1$." Our characterization result shows that this implies $\mu$ is not an efficient iBF. Indeed, it is a not an efficient iBF because it is Pareto dominated by $\mu'$, which is an iBF. In fact, the Pareto-improvement $\mu'$ is obtained by "implementing" this cycle. On $\mu'$, the FIG cycle lets $i_2$ point to (only) $s_2$, $s_2$ point to $i_1$, but does not let $i_1$ point to any school (because $i_1$ is matched to her first choice school). Since $i_3$ does not point to any school (because $i_3$ is matched to her first choice school) either, there is no FIG cycle on $\mu'$. Our characterization result shows that this implies $\mu'$ is an efficient iBF. Indeed, one can verify that there is no iBF that Pareto dominates $\mu'$.

The remainder of this paper proceeds as follows. Section 2 provides a model where we define efficient iBF. Section 3.1 introduces the fair improvement graph (FIG) and FIG cycles. Section 4 provides our main theorem, which characterizes efficient iBF by non-existence of FIG cycles. Section 5 defines the FIG cycles algorithm which outputs an efficient iBF. Section 6 provides various discussions, and Section 7 concludes. Proofs of all results are provided in the Appendix.

## 2. Model

2.1. **Preliminary Definitions.** Let there be a finite set of students $I$ and a finite set of schools $S$. Each student $i$ has a strict preference relation $\succ_i$ over the set of schools and being unmatched (being unmatched is denoted by $\emptyset$). For any $s, s' \in S \cup \{\emptyset\}$, we write $s \succeq_i s'$ if and only if $s \succ_i s'$ or $s = s'$. Each school $s$ has a strict priority order $\succ_s$

[5] In this example, the school with a vacancy points to an unmatched student. The benefit from requiring a school to point to a student matched to another school in the same region does not appear in the current example. We will explain this point in "Example 2, Continued."

over the set of students and leaving a position vacant (which is denoted by $\emptyset$).[6] For any $i, i' \in I \cup \{\emptyset\}$, we write $i \succeq_s i'$ if and only if $i \succ_s i'$ or $i = i'$. Each school $s \in S$ is endowed with a (physical) **capacity** $q_s$, which is a nonnegative integer.

Student $i$ is said to be **acceptable** to school $s$ if $i \succ_s \emptyset$ (and unacceptable otherwise).[7] Similarly, $s$ is acceptable to $i$ if $s \succ_i \emptyset$.[8] It will turn out that only rankings of acceptable partners matter for our analysis, so we often write only acceptable partners to denote preferences and priorities. For example,

$$\succ_i : s, s'$$

means that school $s$ is the most preferred, $s'$ is the second most preferred, and $s$ and $s'$ are the only acceptable schools under preferences $\succ_i$ of student $i$.

A **matching** $\mu$ is a mapping that satisfies (i) $\mu_i \in S \cup \{\emptyset\}$ for all $i \in I$, (ii) $\mu_s \subseteq I$ and $|\mu_s| \leq q_s$ for all $s \in S$, and (iii) for any $i \in I$ and $s \in S$, $\mu_i = s$ if and only if $i \in \mu_s$. That is, a matching simply specifies which student is assigned to which school (if any).

A matching is **individually rational** if no student or school is matched with an unacceptable partner.

Given a matching $\mu$, we say that a student $i$ has **justified envy** to $j \in I$ if there is a school $s \in S$ such that (i) $\mu_j = s$, (ii) $s \succ_i \mu_i$, and (iii) $i \succ_s j$. We say that matching $\mu$ is **fair** if there is no pair of students $(i, j) \in I^2$ such that $i$ has justified envy to $j$.

Finally, a matching $\mu$ **weakly Pareto dominates** a matching $\mu'$ if $\mu_i \succeq_i \mu'_i$ for every $i \in I$. A matching $\mu$ **Pareto dominates** $\mu'$ if $\mu$ weakly Pareto dominates $\mu'$ and $\mu \neq \mu'$.

2.2. **Model of Regions and Efficient iBF.** There is a set of regions, denoted $R$, which is a partition of $I \cup S$. Formally, $R$ satisfies the following conditions.

(1) Each $r \in R$ is a nonempty subset of $I \cup S$.
(2) $r \cap r' = \emptyset$ for any $r, r' \in R$ such that $r \neq r'$.
(3) $\bigcup_{r \in R} r = S \cup I$.

The interpretation is that each $s$ belongs to a single $r \in R$ and each $i$ is a resident of a single $r \in R$. Let $r(s)$ be the region $r$ such that $s \in r$, and similarly for $r(i)$.

We call $\mathcal{E} = (I, S, (\succ_a)_{a \in I \cup S}, (q_s)_{s \in S}, R)$ an **environment**.

We are now ready to introduce the key concept of this paper, "balancedness."

[6]Strictness of priorities is assumed just for the sake of simplicity. In Section 6.3, we consider the case when indifferences are allowed and show that most results carry over to such a case.

[7]We denote singleton set $\{x\}$ by $x$ when there is no confusion.

[8]In some applications, all schools may regard all students as acceptable. None of our results will hinge on the assumption that some students can be unacceptable to some schools.

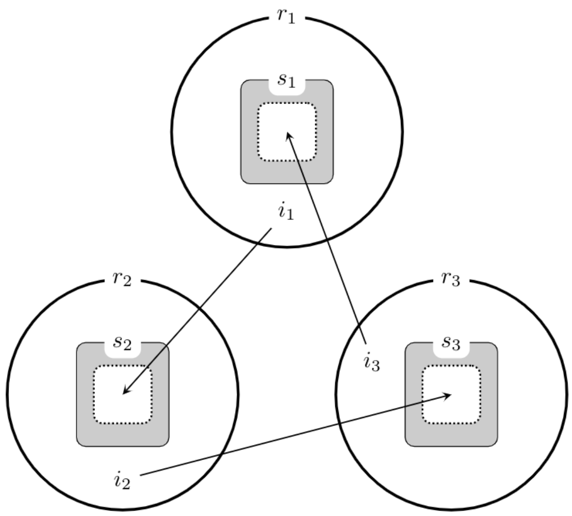


FIGURE 2. Three-way transfer (Example 1). There is one white dotted square in each school, which expresses the fact that the capacity of the school is one.

**Definition 1.** $\mu$ is **balanced** if for each $r \in R$,

$$\underbrace{\sum_{s\in r} |\{i|i \in \mu_s, i \notin r\}|}_{\text{inflow to } r} = \underbrace{\sum_{s\notin r} |\{i|i \in \mu_s, i \in r\}|}_{\text{outflow from } r} \tag{2.1}$$

As eq. (2.1) shows, balancedness means that for any given region $r$, the inflow of students to $r$ is the same as the outflow of students. Note that balancedness is not a "pairwise" notion, that is, it does not necessarily require that for every pair of regions $r$ and $r'$, the number of students who live in $r$ and are matched to a school in $r'$ is the same as the number of students who live in $r'$ and are matched to a school in $r$. The next example illustrates.

**Example 1** (Three-way transfer)**.** Let $I = \{i_1, i_2, i_3\}$, $S = \{s_1, s_2, s_3\}$, $R = \{r_1, r_2, r_3\}$ and $r_k = \{i_k, s_k\}$ for each $k = 1, 2, 3$. Let

$$\mu = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ i_3 & i_1 & i_2 & \emptyset \end{pmatrix}.$$

See Figure 2 for a graphical representation. Note that under $\mu$, the number of students who live in $r_1$ and are matched to a school in $r_2$ is 1 while the number of students who live in $r_2$ and are matched to a school in $r_1$ is 0. Matching $\mu$ is, however, balanced because it satisfies eq. (2.1): For each region, the inflow and outflow are both 1. $\square$

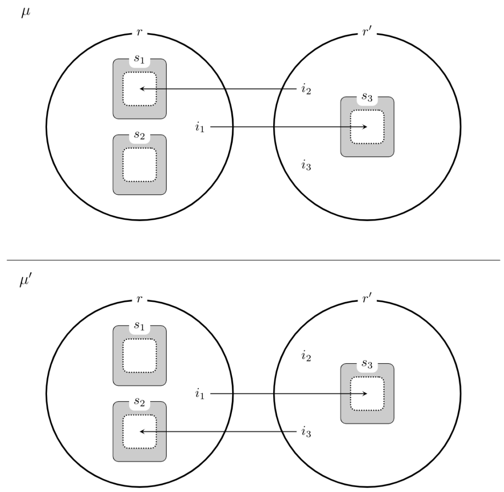


FIGURE 3. Multiple efficient iBFs (Example 2). Both $\mu$ and $\mu'$ are an efficient iBF.

Below is our solution concept in this paper.

**Definition 2.** Matching $\mu$ is said to be a **efficient iBF** if $\mu$ is an iBF and there is no iBF $\mu'$ that Pareto dominates $\mu$.

In the standard environment without the balancedness constraint, there is a unique efficient iBF, which correspond to a "student-optimal stable matching." In our setting, there may be multiple efficient iBFs. The following example illustrates.

**Example 2** (Multiple efficient iBFs)**.** Let $I = \{i_1, i_2, i_3\}$, $S = \{s_1, s_2, s_3\}$, $R = \{r, r'\}$ where $r = \{i_1, s_1, s_2\}$ and $r' = \{i_2, i_3, s_3\}$. Each school has the capacity of one. Student preferences and school priorities are given as follows:

$$\begin{array}{ll} \succ_{i_1}: s_3, s_1, & \succ_{s_1}: i_2, i_1, \\ \succ_{i_2}: s_1, & \succ_{s_2}: i_3, \\ \succ_{i_3}: s_2, & \succ_{s_3}: i_1. \end{array}$$

Let

$$\mu = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ i_2 & \emptyset & i_1, & i_3 \end{pmatrix}, \qquad \mu' = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ \emptyset & s_3 & i_1, & i_2 \end{pmatrix}.$$

See Figure 3 for a graphical representation. We show that both $\mu$ and $\mu'$ are an efficient iBF. To see this, first notice that there are only four individually rational and balanced matchings. They are $\mu$, $\mu'$,

$$\mu'' = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ i_1 & \emptyset & \emptyset, & \emptyset \end{pmatrix},$$

and the empty matching (i.e., the matching where no student is matched to any school). The latter two matchings are Pareto dominated by $\mu$ and $\mu'$ while $\mu$ and $\mu'$ do not Pareto dominate each other, and one can verify by inspection that both $\mu$ and $\mu'$ are fair, and hence an efficient iBF. We note that without the balancedness constraint, the student-optimal stable matching would be:

$$\mu''' = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ i_2 & i_3 & i_1 & \emptyset \end{pmatrix}.$$

But $\mu'''$ is not balanced because the inflow to region $r$ is 2 while the outflow from $r$ is 1. $\square$

## 3. FIG (Fair Improvement Graph) Cycles

3.1. **Definition of FIG Cycle.** The key steps of our analysis involve defining bipartite directed graphs over the sets of students and schools, and identifying cycles on them. A bipartite directed graph on $I$ and $S$, or simply a **graph**, $\mathcal{G} \subseteq (I \times S) \cup (S \times I)$, is a set of ordered pairs of agents in $I \cup S$. An interpretation is that if $(i, s) \in \mathcal{G}$, then there is an arrow pointing from $i$ to $s$. In this case, we say "$i$ points to $s$." The case of $(s, i) \in \mathcal{G}$ is analogous. Given a graph $\mathcal{G}$, a **cycle in** $\mathcal{G}$ is any sequence of the form $(i_1, s_1, i_2, s_2, \ldots, i_m, s_m)$ where

(1) $i_k$ points to $s_k$, i..e, $(i_k, s_k) \in \mathcal{G}$,
(2) $s_k$ points to $i_{k+1}$, i.e, $(s_k, i_{k+1}) \in \mathcal{G}$,
(3) $i_k \neq i_{k'}$ for every $k \neq k'$, and
(4) $s_k \neq s_{k'}$ for every $k \neq k'$,

with $m + 1 = 1$. We will regard any two cycles as defined here, $(i_1, s_1, \ldots, i_m, s_m)$ and $(i_{k+1}, s_{k+1}, i_m, s_m, i_1, s_1, \ldots, i_k, s_k)$, as identical to each other.

Let $D_s^\mu := \{i \in I | s \succ_i \mu_i\}$, and $Top_s(I')$ be the student $i \in I'$ who has the highest priority among those in $I'$ at $\succ_s$.

Now we define a particular type of a graph and a cycle on it. This cycle will be used to characterize efficient iBF as well as to define our algorithm.

**Definition 3.** Given a matching $\mu$, the **fair improvement graph (FIG)** for $\mu$ is a graph such that, for any $i \in I$ and $s \in S$,

(1) student $i \in I$ points to school $s \in S$ if $i = Top_s(D_s^\mu)$ and $i$ is acceptable to $s$, and

(2) school $s \in S$ points to student $i \in I$ if either

(a) $\mu_i = s$, or

(b) $|\mu_s| < q_s$ and, $[i \in r(s)$ and $\mu_i = \emptyset]$ or $\mu_i \in r(s)$.

A **fair improvement graph cycle (FIG cycle)** on $\mu$ is a cycle in the FIG for $\mu$.

In the FIG for a given matching $\mu$, each school can be pointed to by at most one student. A student $i$ can point to a school $s$ when she finds $s$ to be better than her current outcome $\mu_i$, and if she is acceptable to $s$ and the best student for $s$ among those who find $s$ to be an improvement. In this sense, the student that can point to $s$ is selected in the most fair manner, and this is why we call the graph the "fair improvement graph."

On the other hand, a school $s$ can point to a student $i$ in two different cases. The first case is as in other algorithms based on cycles in the literature such as the TTC algorithm (Shapley and Scarf, 1974) and the stable improvement cycles algorithm (Erdil and Ergin, 2008). This is when $i$ is currently matched to $s$, and it is described by item 2a in Definition 3. The second case, described by item 2b in Definition 3, depends on the notion of regions. School $s$ can point to $i'$ if $s$ has a vacancy and either $i$ lives in the region of $s$ and is unmatched, or $i$ is matched to a school in the region. The need for this second case and the logic behind this particular pointing rule will become clear in a later example (Example 5).

Given a matching $\mu$ and a cycle of the form $\mathcal{F} = (i_1, s_1, i_2, s_2, \dots, i_m, s_m)$, call $\mu'$ the **matching generated by** $(\mu, \mathcal{F})$ if

$$\mu'_{i_k} = s_k \text{ for each } k \in \{1, \dots, m\}, \text{ and } \mu'_j = \mu_j \text{ for all } j \in I \setminus \{i_1, \dots, i_m\}.$$

Given a matching $\mu$ and a cycle $\mathcal{F}$, we say that we **implement** $\mathcal{F}$ on $\mu$ when we create the matching generated by $(\mu, \mathcal{F})$.

As we will show in Section 4, whether there exists a FIG cycle is crucial to the characterization of efficient iBF. Also, in the algorithm we define in Section 5, we repeatedly implement FIG cycles. But before stating the formal results that use the notion of FIG cycles, let us illustrate the concept of FIG cycle through a series of examples in the next subsection.

**Remark 1.** Erdil and Ergin (2017) consider a model with weak student preferences (and school priorities) and propose to improve students' welfare by using *chains* in addition to cycles. A chain can start from a matched or unmatched student and ends at a school that has a vacant seat. We could define a chain in our model too, while restricting the pointing from schools to students to the one described in item 2a of Definition 3.

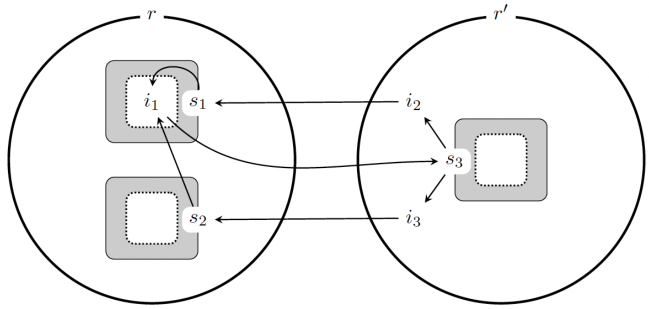


FIGURE 4. Example 2 Continued. There are two FIG cycles.

Let us be forthcoming about the similarity and difference between their chains and our FIG cycles. First, in our model, implementing all chains and cycles under such a pointing rule would violate balancedness (there is no such issue in Erdil and Ergin (2017) as they have no balancedness constraint). For example, in Example 2, there would be one chain going out of $r$ ($i_1 \to s_3$) while there would be two chains going into $r$ ($i_2 \to s_1$ and $i_3 \to s_2$). As we will see in Theorem 1, balancedness is respected if we implement any FIG cycle. Second, one might argue that "connecting" chains might work. That is, we would start from a chain, and at the end of the chain (which is a school), we would find an unmatched student in the same region and see if there was a chain originating from that student. If there was such an arrow, then we would follow the arrows. Continuing this way, if an arrow eventually pointed to a school that had already appeared, then we would call the closed set of arrows a cycle. Cycles constructed in this way turns out to be the same as our FIG cycles. One can view that our pointing rule from schools to students, especially the part described in item 2b of Definition 3, correctly captures how this "connecting" should be done. Third, the reasons behind why there are chains are different. In Erdil and Ergin (2017), a chain is implemented on a stable matching. For the existence of chains it is necessary that the student preferences are weak: if instead the students' preferences are strict, then the last student on the chain would have strictly preferred to be matched to a vacant position in the last school in the chain under the original matching, so the original matching was not stable. In our model, there can exist a chain (defined in the absence of the pointing rule described in item 2b of Definition 3) because of the balancedness constraint. $\square$

### 3.2. **Examples of FIG Cycle.**

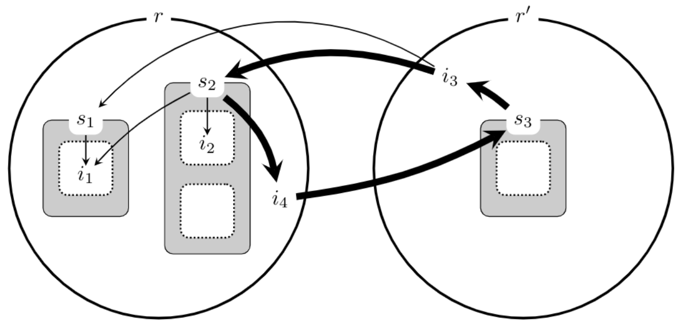


FIGURE 5. Example 3. School $s_2 \in r$ points to an unmatched student $i_4 \in r$ because $s_2$'s capacity is currently not full.

**Example 2, Continued.** Consider the same environment as in Example 2. The FIG for $\mu''$ is drawn in Figure 4. Note that there are two FIG cycles:

$$\mathcal{F} := (i_1, s_3, i_2, s_1) \text{ and } \mathcal{F}' := (i_1, s_3, i_3, s_2).$$

Implementing the former cycle results in $\mu$, and implementing the latter cycle results in $\mu'$. Note that, in this example, it is important that the pointing rule for FIG lets a school point to a student matched to another school in the same region.[9] □

**Example 3** (Capacity matters for FIG cycle)**.** Let $I = \{i_1, i_2, i_3, i_4\}$, $S = \{s_1, s_2, s_3\}$, $R = \{r, r'\}$ where $r = \{i_1, i_2, i_4, s_1, s_2\}$ and $r' = \{i_3, s_3\}$. Schools $s_1$ and $s_3$ have the capacity of one each while $s_2$ has the capacity of two. Student preferences and school priorities are given as follows:

| | |
|---|---|
| $\succ_{i_1}: s_1,$ | $\succ_{s_1}: i_1, i_2, i_3, i_4,$ |
| $\succ_{i_2}: s_2, s_1$ | $\succ_{s_2}: i_2, i_3,$ |
| $\succ_{i_3}: s_1, s_2,$ | $\succ_{s_3}: i_4.$ |
| $\succ_{i_4}: s_1, s_3$ | |

Consider the following matching:

$$\mu = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ i_1 & i_2 & \emptyset & i_3, i_4 \end{pmatrix},$$

which, by inspection, one can show to be fair. This matching as well as the FIG for $\mu$ is drawn in Figure 5. We note two features of this FIG. First, $s_2$ has an arrow directed to an unmatched student $i_4$ even though it is already matched with student $i_2$. This is

[9]This is the point alluded to in footnote 5.

because $s_2$'s capacity is not filled under $\mu$: the capacity is 2 while it is matched to only one student (student $i_2$). Second, there are multiple arrows from student $i_3$. Although the rule for pointing in a FIG implies that each school is pointed to by at most one student, one student can point to multiple schools if she is the "number one" choice from multiple schools. In this example, $i_3$ ranks at the top among all students who prefer $s_1$ to their current match (i.e., among $\{i_3, i_4\}$), and she also ranks at the top among all students who prefer $s_2$ to their current match (i.e., among $\{i_3\}$). Note that there is a single FIG cycle:

$$\mathcal{F} := (i_3, s_2, i_4, s_3).$$

Implementing this cycle, we obtain the following matching:

$$\mu' = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ i_1 & i_2, i_3 & i_4 & \emptyset \end{pmatrix},$$

which is an improvement over $\mu$ and is fair. One can show by inspection that $\mu'$ is also an efficient iBF. □

## 4. Characterization of Efficient iBF

This is the main section of this paper, and we are going to formalize the following claims:

(1) If we implement a FIG cycle on a balanced and fair matching, then it results in a balanced and fair matching that Pareto dominates the original matching (Theorem 1).
(2) If there is no FIC cycle on a given iBF, then that matching is an efficient iBF (Theorem 2).

These results in particular imply the following characterization of efficient iBF: *Given an iBF, it is an efficient iBF if and only if there is no FIG cycle on it* (Corollary 1). In what follows, we will examine each claim and explain their intuition in detail.

**Theorem 1.** *Let $\mu$ be a balanced and fair matching. If there exists a FIG cycle $\mathcal{F}$ on $\mu$, then a matching generated by $(\mu, \mathcal{F})$ is balanced and fair, and Pareto dominates $\mu$. Moreover, if $\mu$ is individually rational, then the matching generated by $(\mu, \mathcal{F})$ is also individually rational.*

An implication of this theorem is that, if we can find a FIG cycle on a given matching, then that matching cannot be an efficient iBF. Thus, one can think of this theorem as providing a necessary condition for a matching to be an efficient iBF. (As we explained,

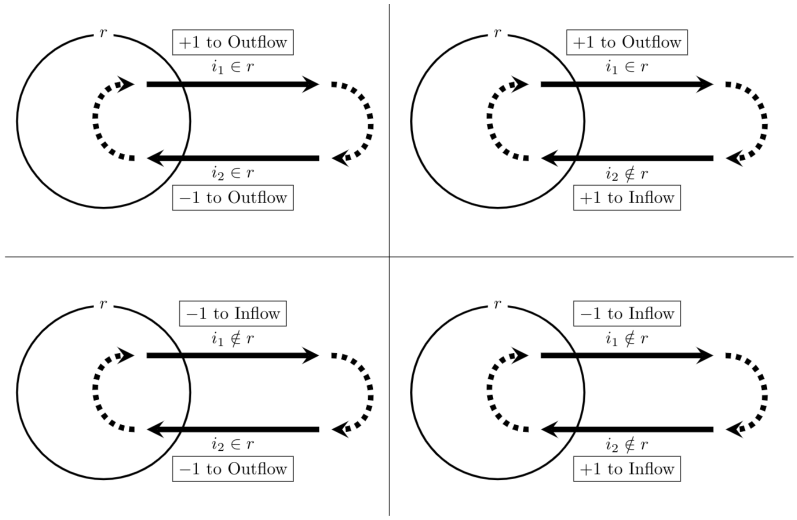


FIGURE 6. Why balancedness is maintained when a cycle is implemented (proof intuition for Theorem 1).

Theorem 2 that we state below provides a sufficient condition, and combining those results will provide a characterization of efficient iBF)

Let us explain the intuition for the proof. For this, let $\mu$ be the original matching and $\mu'$ be the matching generated by $(\mu, \mathcal{F})$ where $\mathcal{F}$ is a FIG cycle on $\mu$. The proof shows Pareto dominance, fairness and balancedness one by one. First, it is straightforward that $\mu'$ Pareto dominates $\mu$.

Second, fairness of $\mu'$ is due to the pointing rule used in the FIG. A crucial step is to show that no student has a justified envy to student $i_k$ who is newly matched to $s_k$ under $\mu'$, where $i_k$ and $s_k$ appear in the FIG cycle (i.e., $i_k$ points to $s_k$ under the FIG).[10] If some student $i$ finds $s_k$ to be better than her match under $\mu'$, then she should have also found $s_k$ to be better than her match under $\mu$ (because $\mu'$ Pareto dominates $\mu$). But the pointing rule for FIG implies that $i_k$ is the best student (according to $s_k$'s priority) among those who found $s_k$ to be better than the match under $\mu$. So, in particular, $i$ is not higher than $i_k$ under $s_k$'s priority. This implies that $i$ cannot have a justified envy to $i_k$ under $\mu'$.

Finally, balancedness of $\mu'$ holds because the FIG cycle is "closed," that is, for any given region $r$, if an arrow from a student goes outside of $r$ along the cycle, then another arrow from a student must come back to $r$ and vice versa, which implies that the number of times the arrows go outside must be equal to the number of times the arrows come back

[10]Justified envy to other students matched to $s_k$ or those that involve other schools can be shown not to exist by using fairness of $\mu$ and the fact that $\mu'$ Pareto dominates $\mu$.

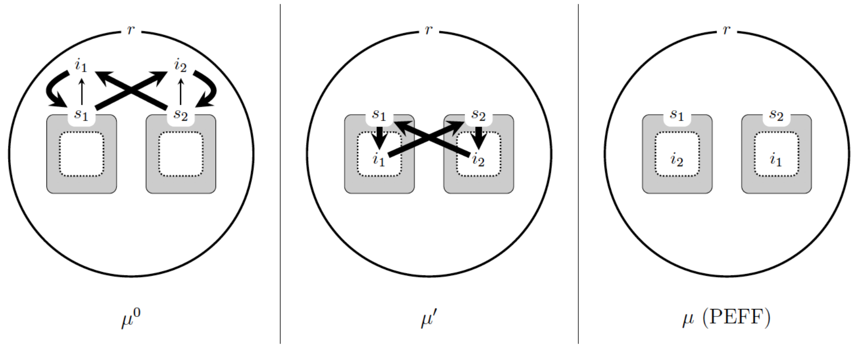


FIGURE 7. Example 4. $\mu$ is a unique efficient iBF, and there is no FIG cycle such that implementing it on $\mu^0$ results in $\mu$. Meanwhile, implementing a FIG cycle on $\mu^0$ (indicated by thick arrows in the left panel of the figure) results in $\mu'$, and implementing a FIG cycle on $\mu'$ (indicated by thick arrows in the middle panel of the figure) results in $\mu$.

to $r$. Whether balancedness is maintained by implementing a cycle may not be obvious due to the fact that an outgoing arrow may carry a student who lives in $r$ or one who does not live in $r$, and similarly an incoming arrow may carry either type of a student. Figure 6 shows that in every possible case, balancedness is maintained when a cycle is implemented.[11] Feasibility would also be violated if an arrow that goes across regions can originate from a school, a possibility that the definition of FIG prohibits. Example 5 illustrates this point.

Note that Theorem 1 does not assert that implementing a FIG cycle on a balanced and fair matching necessarily results in an efficient iBF. Indeed, the following example presents a case in which one needs to implement FIG cycles more than once to reach an efficient iBF.

**Example 4** (Multiple rounds of FIG cycles)**.** Let $I = \{i_1, i_2\}$, $S = \{s_1, s_2\}$, $R = \{r\}$ . Each school has the capacity of one. Student preferences and school priorities are given

[11]Hafalir, Kojima and Yenmez (2018) study TTC under a variety of constraints, one of which is the balancedness constraints of the present paper. They verify that the balancedness constraints satisfy a condition called M-concavity, which Suzuki, Tamura and Yokoo (2018) showed is sufficient for the outcome of a certain version of TTC to satisfy balancedness. Although their TTC algorithm is substantially different from ours, one might also be able to use a similar indirect approach to establish balancedness of the outcome of our algorithm.

as follows:

$$\succ_{i_1}\colon s_2, s_1, \qquad\qquad \succ_{s_1}\colon i_1, i_2,$$
$$\succ_{i_2}\colon s_1, s_2, \qquad\qquad \succ_{s_2}\colon i_2, i_1.$$

Let $\mu^0$ be the empty matching. Starting from $\mu^0$, there are three FIG cycles:

$$\mathcal{F} := (i_1, s_1),\ \mathcal{F}' := (i_2, s_2),\ \text{ and }\ \mathcal{F}'' := (i_1, s_1, i_2, s_2).$$

It is straightforward to see that there is a unique efficient iBF, which is

$$\mu = \begin{pmatrix} s_1 & s_2 & \emptyset \\ i_2 & i_1 & \emptyset \end{pmatrix}.$$

However, $\mu$ is not a matching generated by $(\mu^0, \tilde{\mathcal{F}})$ where $\tilde{\mathcal{F}}$ is either $\mathcal{F}$, $\mathcal{F}'$ or $\mathcal{F}''$. The reason is that, for example, student $i_1$ does not point to $s_2$ in the FIG for $\mu^0$ because she is not the top choice of $s_2$ among those who point to $s_2$ ($s_2$'s top choice is $i_2$).

The efficient iBF can be obtained by "two rounds of FIG cycles" starting from $\mu^0$: First, implementing the FIG cycle $\mathcal{F}''$ on $\mu^0$, we obtain:

$$\mu' = \begin{pmatrix} s_1 & s_2 & \emptyset \\ i_1 & i_2 & \emptyset \end{pmatrix}.$$

Second, implementing

$$\mathcal{F}''' := (i_1, s_2, i_2, s_1)$$

on $\mu'$ results in $\mu$. $\square$

Another comment is in order. Theorem 1 asserts that, among other things, implementing a FIG cycle on a balanced matching results in a balanced matching. This property depends on a somewhat subtle manner in which we define the FIG, as the following example illustrates.

**Example 5** (Pointing rule for FIG)**.** Let $I = \{i_1, i_2\}$, $S = \{s_1, s_2, s_3\}$, $R = \{r, r'\}$ where $r = \{i_1, s_1, s_2\}$ and $r' = \{i_2, s_3\}$. Each school has the capacity of one. Student preferences and school priorities are given as follows:

$$\succ_{i_1}\colon s_2, s_3, s_1 \qquad\qquad \succ_{s_1}\colon i_1, i_2$$
$$\succ_{i_2}\colon s_1, s_2 \qquad\qquad \succ_{s_2}\colon i_2, i_1,$$
$$\succ_{s_3}\colon i_1.$$

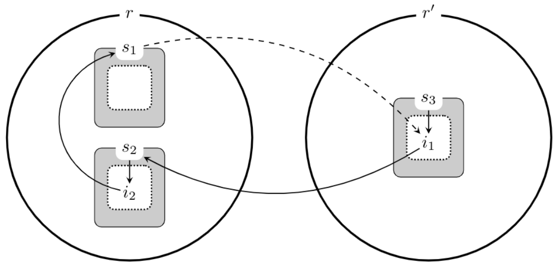


FIGURE 8. Example 5. Student $i_1$ lives in $r$ and student $i_2$ lives in $r'$. If we allowed $s_1$ to point to $i_1$ (the dashed arrow) in the FIG and implemented the resulting cycle $(i_2, s_1, i_1, s_2)$, the balancedness constraint would be violated.

Consider the following matching:

$$\mu = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ \emptyset & i_2 & i_1 & \emptyset \end{pmatrix}.$$

Note that this matching is a matching generated by $(\mu^0, \mathcal{F})$ where $\mu^0$ denotes the empty matching and we let $\mathcal{F} := (i_1, s_3, i_2, s_2)$. The cycle $\mathcal{F}$ is a FIG cycle and $\mu$ is fair.

Now, consider the FIG on $\mu$, which is depicted in Figure 8. Suppose that we hypothetically allow each school with a vacancy to point to *any* student that lives in the same region, irrespective of where that student is currently matched. This means that $s_1$ points to $i_1$ (which is *not* allowed in the FIG, according its pointing rule). This additional pointing is depicted by a dotted arrow in Figure 8. With this arrow, there is a cycle $\mathcal{F}' := (i_1, s_2, i_2, s_1)$. However, the matching generated by $(\mu, \mathcal{F}')$ violates the balancedness constraint. The reason is that implementing this cycle on $\mu$ would result in one less outflow for region $r$ while the inflow does not change. This happens because the arrow going out of region $r$ starts from a school, not a student. In the definition of the FIG, all arrows going out of regions are from students. As we explained when providing the intuition for the proof of balancedness in Theorem 1 (see Figure 6), this feature ensures that implementing any FIG cycle on any balanced matching results in a balanced matching. $\square$

**Theorem 2.** *Let $\mu$ be an iBF. $\mu$ is an efficient iBF if there exists no FIG cycle on $\mu$.*

The proof is by contraposition. That is, we take $\mu$ that is an iBF and assume that there is another iBF $\mu'$ that Pareto dominates $\mu$. Then we show that there exists a FIG cycle on $\mu$.

To find a FIG cycle, we construct a graph. In this graph, the students who are associated with arrows are those whose outcomes are different between $\mu$ and $\mu'$. We denote by $I'$ the set of those students. Meanwhile, the schools with arrows are the ones that are matched to students in $I'$ under $\mu'$ (which means that these are the schools that are matched to some new students under $\mu'$ relative to $\mu$). We denote by $S(\mu')$ the set of those schools. Also, let $S(\mu)$ be the set of schools that students in $I'$ are matched with under $\mu$.

To understand the complication in the proof, consider first a simple case: there is a single region $r$ (i.e., a standard environment without a balancedness constraint), and $\mu$ is a stable matching (in the standard sense). Then, since stability implies non-wastefulness, it is immediate that $S(\mu') = S(\mu)$. We can have each school $s$ point to students in $\mu_s \cap I'$, and allow a student in $I'$ point to $s$ if she is the top student among $I'$ who regard $s$ to be an improvement relative to $\mu$. This way, each school is pointed to by one student, and each student is pointed to by one school. Thus, there is a cycle, and with some work, one can prove that such a cycle must be a FIG cycle. This is essentially the same method as what is used in Erdil and Ergin (2008).[12]

In our problem, complication arises for two reasons. First, fairness alone does not imply non-wastefulness, so $\mu$ may have some waste (for example, consider the empty matching $\mu^0$ in Example 4). This means that $S(\mu')$ may not be equal to $S(\mu)$, so a graph constructed in the above manner might not have a cycle (schools in $S(\mu') \setminus S(\mu)$ would have no outgoing arrow, and those in $S(\mu) \setminus S(\mu')$ would have no incoming arrow). This suggests we need an alternative way of forming a graph. Second, arrows might go in and out of any region, hence in defining the alternative graph, we must make sure that the balancedness constraint would be respected when implementing a cycle in the graph.

We overcome these difficulties by constructing a graph, denoted $\mathcal{G}(\mu, \mu')$, in the following manner. First, each school in $S(\mu')$ is pointed to by the top student among those in $I'$ who regard the school to be an improvement relative to $\mu$, just as in the "simple case" above. Second, each student $i$ in $I'$ is pointed to by a school in the following three different ways depending on $i$'s outcome under $\mu$:

(1) If $\mu_i \in S(\mu')$, then we let $\mu_i$ point to $i$. This is the case that is analogous to the "simple case" explained above.

[12]We note that Erdil and Ergin (2008) allow school priorities to be weak, while we assume strict preferences here. However, as explained in Section 6.3, our analysis extends to the case with weak priorities without any significant change.

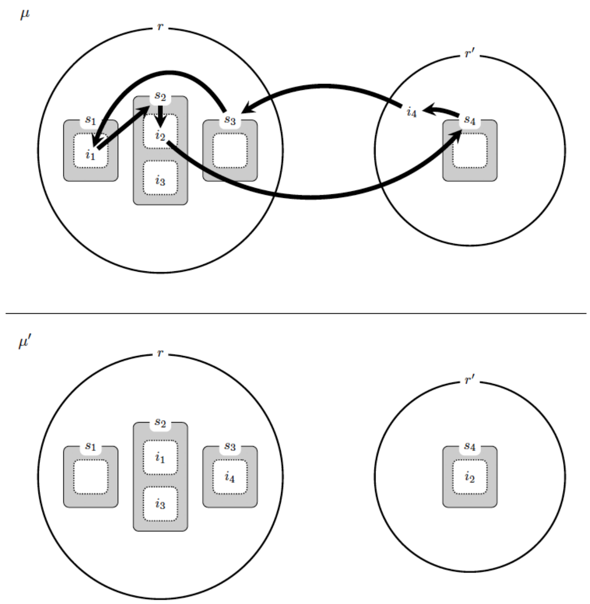


FIGURE 9. Example 6. Construction of a FIG cycle in the proof of Theorem 2.

(2) If $\mu_i \in S \setminus S(\mu')$, then we can show, using individual rationality and balancedness, that we can find a school that resides in $\mu_i$'s region and belongs to $S(\mu')$ whose capacity is not filled under $\mu$.[13] We let such a school point to $i$.

(3) If $\mu_i = \emptyset$, then we can show, again using individual rationality and balancedness, that we can find a school that resides in $i$'s region and belongs to $S(\mu')$ whose capacity is not filled under $\mu$. We let such a school point to $i$.

Since each school is pointed to by one student, and each student is pointed to by one school, $\mathcal{G}(\mu, \mu')$ has a cycle. The proof shows that any cycle in this graph is a FIG cycle.

The next example illustrates how our construction works in a specific environment.

**Example 6** (Construction of a FIG cycle)**.** Let $I = \{i_1, i_2, i_3, i_4\}$, $S = \{s_1, s_2, s_3, s_4\}$, $R = \{r, r'\}$ where $r = \{i_1, i_2, i_3, s_1, s_2, s_3\}$ and $r' = \{i_4, s_4\}$. School $s_2$ has the capacity of two while all other schools have the capacity of one. Student preferences and school

[13]To show the existence of such a school, the proof constructs another graph, which is so much fun, but unfortunately we have to omit the detail for the space consideration. See the proof in the Appendix for the full experience of the proof.

priorities are given as follows:

$$\succ_{i_1}: s_2, s_1, \qquad \succ_{s_1}: i_1,$$
$$\succ_{i_2}: s_4, s_2, \qquad \succ_{s_2}: i_2, i_3, i_1, i_4,$$
$$\succ_{i_3}: s_2, \qquad \succ_{s_3}: i_4,$$
$$\succ_{i_4}: s_2, s_3, \qquad \succ_{s_4}: i_2.$$

Consider the following matchings:

$$\mu = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & \emptyset \\ i_1 & i_2, i_3 & \emptyset & \emptyset & i_4 \end{pmatrix}, \qquad \mu' = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & \emptyset \\ \emptyset & i_1, i_3 & i_4, & i_2 & \emptyset \end{pmatrix}.$$

They are both an iBF, and $\mu'$ Pareto dominates $\mu$.

Let us explain how to obtain the graph $\mathcal{G}(\mu, \mu')$ and a FIG cycle in this environment (see Figure 9 for a graphical illustration). First, note that $I' = \{i_1, i_2, i_4\}$, $S(\mu) = \{s_1, s_2\}$, and $S(\mu') = \{s_2, s_3, s_4\}$. Recall that a student in $I'$ points to a school in $S(\mu')$ if and only if she is the top student according to the school's priority among those who prefer the school to their current match. Hence, the edges originating at a student in $\mathcal{G}(\mu, \mu')$ are $(i_1, s_2), (i_2, s_4)$, and $(i_4, s_3)$. Next, we illustrate how a student is pointed to by a school.

(1) $\mu_{i_1} = s_1 \in S \setminus S(\mu')$. This is case 2 of the aforementioned pointing rule from a school to a student. School $s_3$ is the only school that resides in $\mu_{i_1}$'s region $r$ and belongs to $S(\mu')$ whose capacity is not filled under $\mu$. Thus, (only) $s_3$ points to $i_1$.

(2) $\mu_{i_2} = s_2 \in S(\mu')$. This is case 1 of the pointing rule, and thus $s_2$ points to $i_2$.

(3) $\mu_{i_4} = \emptyset$. This is case 3 of the pointing rule. School $s_4$ is the only school that resides in $i_4$'s region $r'$ and belongs to $S(\mu')$ whose capacity is not filled under $\mu$. Thus, (only) $s_4$ points to $i_4$.

Overall, the graph $\mathcal{G}(\mu, \mu')$ can be drawn as in Figure 9. There is a unique cycle on $G(\mu, \mu')$, which is $(i_1, s_2, i_2, s_4, i_4, s_3)$. One can check that this is a FIG cycle as well. □

Theorem 1 and Theorem 2 together imply the following characterization of efficient iBF

**Corollary 1.** *Let $\mu$ be an iBF. $\mu$ is an efficient iBF if and only if there exists no FIG cycle on $\mu$.*

## 5. FIG Cycles Algorithm

Let $\tilde{\mu}$ be an arbitrary iBF, e.g., the empty matching. Building on Corollary 1, we define a **FIG cycles algorithm** on $\tilde{\mu}$ as follows.

**FIG Cycles Algorithm:**

Step 0: Let $\tilde{\mu}$ be an arbitrary iBF, e.g., the empty matching. Let $\mu^0 = \tilde{\mu}$ and move to Step 1.

Step $l$ ($l \geq 1$): If there is no FIG cycle on $\mu^{l-1}$, terminate the algorithm and output $\mu^{l-1}$. Otherwise, choose a FIG cycle $\mathcal{F}$ on $\mu^{l-1}$ arbitrarily and let $\mu^l$ be the matching generated by $(\mu^{l-1}, \mathcal{F})$, and go to Step $l+1$.

**Corollary 2.** *Let $\tilde{\mu}$ be an arbitrary iBF. The FIG cycles algorithm on $\tilde{\mu}$ runs in a polynomial time, and its output is an efficient iBF.*

*Proof.* Take an arbitrary iBF and denote it by $\tilde{\mu}$. Theorems 1 and 2 together show that the output of the FIG cycles algorithm on $\tilde{\mu}$ is an efficient iBF. To show that the algorithm runs in a polynomial time, first note that each student can only become better off while running the algorithm, and at least one student must be made strictly better off at each step as long as the algorithm does not terminate in that step. Therefore, at most $|I| \times |S|$ steps are necessary for terminating the algorithm. Second, within each step, finding a FIG cycle can be done in a polynomial time.[14] These two observations show that the algorithm runs in a polynomial time, as desired. $\square$

In some applications, a school may give a higher priority to residents of that school's region than non-residents. Let us now consider such a case. Formally, assume that $i \succ_s j$ for each $r \in R$, $s \in r \cap S$, $i \in r \cap I$, and $j \in I \setminus r$.

We consider a special fair matching, namely the matching that is produced by running the standard deferred acceptance algorithm of Gale and Shapley (1962), separately in each region. Formally, it is defined as follows.

**Definition 4.** A **region-wise student-optimal stable matching** $\mu^{RW}$ is a matching that satisfies the following:

(1) For each $r \in R$ and each $i \in r$, we have $\mu_i^{RW} \in (r \cap S) \cup \{\emptyset\}$.
(2) $\mu^{RW}$ is individually rational.
(3) For each $r \in R$, there is no pair of a student and a school $i, s \in r$ such that $s \succ_i \mu_i^{RW}$, $i \succ_s \emptyset$, and $|\mu_s^{RW}| < q_s$.
(4) For each $r \in R$, there is no pair of students $i, i' \in r$ such that $i$ has justified envy to $i'$ under $\mu^{RW}$.
(5) $\mu^{RW}$ weakly Pareto dominates all matchings that satisfy conditions (1), (2), (3) and (4).

[14]There are well-known polynomial-time algorithms that identify a cycle if one exists and otherwise show that there is no cycle. The "depth-first search" algorithm, for example, has the running time of $O(|I| \times |S|)$ (Cormen et al., 2001).

Intuitively, a region-wise student-optimal stable matching requires that a matching restricted to each region (i.e., consider the students and schools that reside in that region) is a student-optimal stable matching in the standard sense. By Gale and Shapley (1962), a region-wise student-optimal stable matching always exists, and it can be obtained by their deferred acceptance algorithm in a polynomial time. By the assumption in this section that schools rank students from the same region higher than those from other regions (so a student living in $r$ cannot have justified envy to another student living in $r'$ if the latter student is matched to a school in $r'$), it is fair.[15]

The next corollary considers the case when we run a FIG cycles algorithm starting from this matching.

**Corollary 3.** *The output of a FIG cycles algorithm starting from the region-wise student-optimal stable matching $\mu^{RW}$ is an efficient iBF and weakly Pareto dominates $\mu^{RW}$.*

*Proof.* Let $\mu^{RW}$ be the region-wise student-optimal stable matching. Note that individual rationality and balancedness of $\mu^{RW}$ follow from conditions (2) and (1) of Definition 4, respectively. $\mu^{RW}$ is also fair as we have explained after providing the statement of Definition 4 (we use condition (4) and our assumption on the priorities). Hence, Corollary 2 implies that the output of a FIG cycles algorithm starting from $\mu^{RW}$ is an efficient iBF.

Finally, Theorem 1 and the definition of FIG cycles algorithm imply that the output Pareto dominates $\mu^{RW}$. □

## 6. Discussions

This section provides a number of discussions. Section 6.1 examines the strategy properties. We show that our mechanism based on the FIG cycle algorithm turns out not to be strategy-proof. We, however, also show that there is no strategy-proof mechanism that always outputs an efficient iBF. In Section 6.2, we provide comparative statics to evaluate the effect of merging and splitting regions. Section 6.3 considers the case with weak priority, which often arises in school choice applications. In Section 6.4, we review the related literature.

---

[15]We wrote Definition 4 following the most standard way, but some conditions and qualifiers are redundant because condition (5) holds and, as we explained, $\mu^{RW}$ is fair. A simplified (equivalent) definition requires the following three conditions:

(1) For each $r \in R$ and each $i \in r$, we have $\mu_i^{RW} \in (r \cap S) \cup \{\emptyset\}$.

(4') $\mu^{RW}$ is fair.

(5') $\mu^{RW}$ weakly Pareto dominates all matchings that satisfy conditions (1) and (4').

6.1. **Strategic Property.** A mechanism $\varphi$ is a function from the set of student preference profiles to the set of matchings. Mechanism $\varphi$ is **strategy-proof** if

$$\varphi_i(\succ) \succeq_i \varphi_i(\succ'_i, \succ_{-i}),$$

for every student preference profile $\succ$, $i \in I$, and student preference $\succ'_i$.

**Theorem 3.** *There exists no strategy-proof mechanism that outputs an efficient iBF for all preference profiles.*

*Proof.* We prove the result by presenting an example. Let $I = \{i_1, i_2, i_3\}$ and $S = \{s_1, s_2, s_3\}$. Let there be two regions, $r = \{s_1, i_1, i_2\}$ and $r' = \{s_2, s_3, i_3\}$. Each school has the capacity of one. Student preferences and school priorities are given as follows:

$$\begin{aligned} &\succ_{i_1}: s_2, && \succ_{s_1}: i_1, i_2, i_3, \\ &\succ_{i_2}: s_3, && \succ_{s_2}: i_3, i_2, i_1, \\ &\succ_{i_3}: s_1, && \succ_{s_3}: i_3, i_1, i_2. \end{aligned}$$

In this environment, there are two efficient iBFs:

$$\mu = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ i_3 & i_1 & \emptyset, & i_2 \end{pmatrix}, \qquad \mu' = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ i_3 & \emptyset & i_2, & i_1 \end{pmatrix}.$$

Fix a mechanism that outputs an efficient iBF for all preference profiles, $\varphi$. It must be either $\varphi(\succ) = \mu$ or $\varphi(\succ) = \mu'$.

Suppose $\varphi(\succ) = \mu$. Then, consider $\succ'_{i_2}: s_3, s_2$. The unique efficient iBF at $\succ' := (\succ'_{i_2}, \succ_{-i_2})$ is $\mu'$, so $\varphi(\succ') = \mu'$. Noting that $\mu'_{i_2} = s_3 \succ_{i_2} \emptyset = \mu_{i_2}$, we have obtained that $\varphi_{i_2}(\succ) \not\succeq_{i_2} \varphi_{i_2}(\succ'_{i_2}, \succ_{-i_2})$, violating strategy-proofness.

Next, suppose $\varphi(\succ) = \mu'$. Then, by considering $\succ'_{i_1}: s_2, s_3$ and following a symmetric argument, we conclude that $\varphi$ is not strategy-proof. This completes the proof. $\square$

6.2. **Comparative Statics.** We say that an environment $\mathcal{E} = (I, S, (\succ_a)_{a \in I \cup S}, (q_s)_{s \in S}, R)$ **is a result of mergers from** another environment $\mathcal{E}' = (I', S', (\succ'_a)_{a \in I' \cup S'}, (q'_s)_{s \in S'}, R')$ if $I = I', S = S'$, $\succ_a = \succ'_a$ for every $a \in I \cup S$, $q_s = q'_s$ for every $s \in S$ and, for each $r \in R$, $r$ is a union of (possibly one) regions of $R'$. That is, some regions in $\mathcal{E}'$ merge to form a region in $\mathcal{E}$, but otherwise all the primitives are unchanged between the two environments.

**Proposition 1.** *Suppose that $\mathcal{E}$ is a result of mergers from $\mathcal{E}'$. Then, for any matching $\mu'$ that is an efficient iBF at $\mathcal{E}'$, there exists a matching $\mu$ that is an efficient iBF at $\mathcal{E}$ such that $\mu$ weakly Pareto dominates $\mu'$.*

*Proof.* If $\mu'$ is an efficient iBF at $\mathcal{E}$, then the conclusion of the proposition holds by setting $\mu := \mu'$. Thus, suppose that $\mu'$ is not an efficient iBF at $\mathcal{E}$. Note that $\mu'$ is balanced and fair at $\mathcal{E}$. Then, run an FIC algorithm in which the initial matching is $\mu'$. Let $\mu$ be the outcome of the FIC algorithm, which is guaranteed to stop in a finite number of steps. By Theorem 2, $\mu$ is an efficient iBF at $\mathcal{E}$. Moreover, by Theorem 1, $\mu$ Pareto dominates $\mu'$, concluding the proof. □

In the following example, merging regions strictly Pareto-improves the outcomes for students.

**Example 7** (An instance in which merging regions makes all students better off)**.** Let $I = \{i\}$ and $S = \{s\}$. Let there be two regions, $r = \{i\}$, and $r' = \{s\}$.[16] School $s$ has the capacity of one. Student $i$ finds school $s$ to be acceptable and school $s$ finds student $i$ to be acceptable. In this environment, there is a unique efficient iBF:

$$\mu = \begin{pmatrix} s & \emptyset \\ \emptyset & i \end{pmatrix}.$$

If regions $r_1$ and $r_2$ are merged, then there is a unique efficient iBF:

$$\mu' = \begin{pmatrix} s & \emptyset \\ i & \emptyset \end{pmatrix}.$$

Since $\mu'_i \succ \mu_i$ and $i$ is the only student in the market, this means that merging the regions Pareto-improved all the students in this example. The intuition is simple: The merge reduced the constraint of balanced exchange between regions $r$ and $r'$, so after the merger, the student $i$ can go to school $s$.

The following example shows that a kind of converse of Proposition 1 does not hold.

**Example 8** (An instance in which splitting a region inevitably makes some student better off)**.** Let $I = \{i_1, i_2\}$ and $S = \{s_1, s_2, s_3\}$. Let there be two regions, $r = \{s_1, s_2, i_1\}$ and $r' = \{s_3, i_2\}$. Each school has the capacity of one. Student preferences and school priorities are given as follows:

$$\succ_{i_1}\colon s_2, s_3, \qquad \succ_{s_1}\colon i_2,$$

$$\succ_{i_2}\colon s_1, \qquad \succ_{s_2}\colon i_1,$$

$$\succ_{s_3}\colon i_1.$$

[16]In this example, region $r$ does not have a school and region $r'$ does not have a student. These features are not necessary to make our point.

In this environment, there is an efficient iBF:

$$\mu = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ \emptyset & i_1 & \emptyset & i_2 \end{pmatrix}.$$

If region $r$ is split into two regions, $r_1 = \{s_1, i_1\}$ and $r_2 = \{s_2\}$, then there is a unique efficient iBF:

$$\mu' = \begin{pmatrix} s_1 & s_2 & s_3 & \emptyset \\ i_2 & \emptyset & i_1 & \emptyset \end{pmatrix}.$$

Note that $i_2$ is better off as a result of the split. The intuition is the following: Before the split, $i_2$ in $r'$ was unable to match with $s_1$ in $r$ as there was no student who wanted to come from $r$ to his region $r'$. However, the split made it impossible for $i_1$ in $r_1$ (which was part of $r$) to go to $s_2$, and she is now interested in coming to $r'$. This made it possible to implement a swap between $r_1$ and $r'$.

6.3. **Weak Priorities.** In applications such as daycare allocations and school choice, schools are sometimes endowed with weak priorities. In fact, Erdil and Ergin (2008) consider weak priorities and propose an algorithm based on cycles to improve upon the deferred acceptance algorithm with tie-breaking, albeit in a setting without our balancedness constraints. Accordingly, a natural question would be whether our analysis extends to cases where priorities are allowed to be weak in the presence of balancedness constraints. As it turns out, all of our results go through. In particular, the conclusions of Theorems 1 and 2 hold without any change.[17]

The only difference from the case of strict priorities is the following. We stated that, in the FIG or $\mathcal{G}(\mu, \mu')$ (a graph which appears in the proof of Theorem 2), each school can be pointed to by at most one student. This was because a student can point to a school only if she is the top student among those who regard the school as an improvement. Under strict priorities, there is a unique "top" student, which was why each school can be pointed to by at most one student. Under weak priorities, however, there can be multiple "top" students, so a school can be pointed to by multiple students. But this change does not affect the proof.

[17]Under weak priority, we say that a student is acceptable to a school if she is ranked *weakly* higher than the outside option, and modify the definitions of individual rationality, efficient iBF and FIG by adopting this definition of acceptability in the relevant parts of those concepts. The proofs change accordingly in a straightforward manner.

6.4. **Related Literature.** The aim of our study is to improve efficiency in situations in which matching is organized at a small local level. Sharing this interest with the present paper, Hafalir, Kojima and Yenmez (2018) study a variety of constraints in the context of interdistrict school choice, and one of the constraints studied there is the balancedness constraint. Differently from our analysis, they study conditions under which the outcome of the standard deferred acceptance algorithm satisfies the constraint, which turns out to be very restrictive. The present paper offers a way to improve upon matching mechanisms organized separately in each district even if the deferred acceptance algorithm does not produce a balanced matching. To our knowledge, the balancedness condition was introduced and studied first by Dur and Ünver (2019), although the setting of tuition exchange they study is different from ours or Hafalir, Kojima and Yenmez (2018) in several important respects, making their analysis and ours logically unrelated.

Our algorithm is based on a number of Pareto-improving cycles among students and schools. At a high level, this is a common idea and is shared by many other algorithms, including Gale's celebrated TTC algorithm in Shapley and Scarf (1974). The difference of our algorithm from TTC is that we construct cycles in a more subtle and nuanced manner, taking school priorities into account in particular, so that implementing the cycles will keep fairness of the original matching. Closer to our algorithm are those due to Erdil and Ergin (2008, 2017) who, like us, provide iterative algorithms that improve efficiency while retaining fairness.[18] Similarities and differences between those studies and ours are illustrated in detail in Sections 1.1 and 4 and remark 1.

Lastly, this paper is part of the growing literature on matching theory and market design. Ever since the seminal contribution by Gale and Shapley (1962), matching theory proved to be a source of fruitful insights. What is especially remarkable is its use in applications to market design. Research in this field has been successfully applied to various problems such as medical match (Roth, 1984; Roth and Peranson, 1999), school choice (Balinski and Sönmez, 1999; Abdulkadiroğlu and Sönmez, 2003), organ donation (Roth, Sönmez and Ünver, 2004, 2005, 2007), and course allocation (Sönmez and Ünver, 2010; Budish and Cantillon, 2012), among others.

## 7. Conclusion

This paper studied a school-choice matching model that allows for inter-district transfer of students. Given any matching, we defined a directed bipartite graph (the "FIG")

[18]Algorithms based on analogous ideas have been adapted to other settings by Combe, Tercieux and Terrier (2018) and Erdil and Kumano (2019), among others.

in which the nodes represent students and schools while the edges are constructed using student preferences, school priorities, and the information about the current matching. Using this graph, we characterized the set of efficient iBFs (individually rational, balanced, and fair matchings) by non-existence of a cycle that would improve the welfare of the students involved in the cycle. This led us to define the FIG cycles algorithm that computes an efficient iBF in polynomial time. We analyzed various additional issues, such as comparative statics where regions are merged or split.

Our results provide a concrete solution to the problem of improving the welfare when matching is organized at small local levels. Such problems are abundant in real markets: for example, in the City of Tokyo, each of the 23 municipalities independently organizes a daycare matching market. Our method does not require these 23 municipalities to fully integrate with one another so that any transfers of students are allowed between different municipalities, which we view would be an unrealistic solution. We, by contrast, only require partial integration where transfer can be made to the extent that the balancedness condition is respected. We view such a compromise to be a realistic solution and hope our method is used in practice.

## References

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

**Agarwal, Nikhil, Itai Ashlagi, Eduardo Azevedo, Clayton R Featherstone, and Ömer Karaduman.** 2019. "Market failure in kidney exchange." *American Economic Review*, 109(11): 4026–70.

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

**Budish, E., and E. Cantillon.** 2012. "The Multi-unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard." *American Economic Review*, 102: 2237–2271.

**Combe, Julien, Olivier Tercieux, and Camille Terrier.** 2018. "The design of teacher assignment: Theory and evidence." *Unpublished paper, University College London.[1310].*

**Cormen, Thomas H, Charles E Leiserson, Ronald L Rivest, and Clifford Stein.** 2001. *Introduction to Algorithms.* MIT Press and McGraw-Hill.

**Dur, Umut Mert, and M Utku Ünver.** 2019. "Two-sided matching via balanced exchange." *Journal of Political Economy*, 127(3): 1156–1177.

**Education Commission of the States.** 2017. “Open Enrollment 50-State Report - All Data Points.” https://ecs.secure.force.com/mbdata/mbquest4e?rep=OE1705.

**Erdil, Aytek, and Haluk Ergin.** 2008. “What’s the Matter with Tie-Breaking? Improving Efficiency in School Choice.” *American Economic Review*, 98: 669–689.

**Erdil, Aytek, and Haluk Ergin.** 2017. “Two-sided matching with indifferences.” *Journal of Economic Theory*, 171: 268–292.

**Erdil, Aytek, and Taro Kumano.** 2019. “Efficiency and stability under substitutable priorities with ties.” *Journal of Economic Theory*, 184: 104950.

**Gale, David, and Lloyd S. Shapley.** 1962. “College Admissions and the Stability of Marriage.” *American Mathematical Monthly*, 69: 9–15.

**Hafalir, Isa Emin, Fuhito Kojima, and M Bumin Yenmez.** 2018. “Interdistrict school choice: A theory of student assignment.” *Available at SSRN 3307731*.

**Roth, Alvin E.** 1984. “The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory.” *Journal of Political Economy*, 92: 991–1016.

**Roth, Alvin E., and Elliott Peranson.** 1999. “The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design.” *American Economic Review*, 89: 748–780.

**Roth, Alvin E., Tayfun Sönmez, and Utku Ünver.** 2004. “Kidney Exchange.” *Quarterly Journal of Economics*, 119: 457–488.

**Roth, Alvin E., Tayfun Sönmez, and Utku Ünver.** 2005. “Pairwise Kidney Exchange.” *Journal of Economic Theory*, 125: 151–188.

**Roth, Alvin E., Tayfun Sönmez, and Utku Ünver.** 2007. “Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences.” *American Economic Review*, 97: 828–851.

**Shapley, Lloyd, and Herbert Scarf.** 1974. “On cores and indivisibility.” *Journal of mathematical economics*, 1(1): 23–37.

**Sönmez, Tayfun, and Utku Ünver.** 2010. “Course Bidding at Business Schools.” *International Economic Review*, 51(1): 99–123.

**Suzuki, Takamasa, Akihisa Tamura, and Makoto Yokoo.** 2018. “Efficient allocation mechanism with endowments and distributional constraints.” 50–58.

# Appendix A. Proofs

## A.1. **Proof of Theorem 1.**

*Proof.* Fix $\mu$ and a FIG cycle $\mathcal{F} = (i_1, s_1, i_2, s_2, \ldots, i_m, s_m)$ on $\mu$. Let $\mu'$ be the matching that is generated by $(\mu, \mathcal{F})$. Clearly, $\mu'$ Pareto dominates $\mu$.

To show that $\mu'$ is fair, notice that $\mu'_i \succeq_i \mu_i$ for all $i \in I$ by the definition of FIG cycle. This implies that, for every $i \notin \{i_1, \ldots, i_m\}$, no one has justified envy to $i$ under $\mu'$ because $\mu$ is fair. Thus, it remains to show that for each $k \in \{1, \ldots, m\}$, no one has justified envy to $i_k$ under $\mu'$. To see this, note that, again by $\mu'_i \succeq_i \mu_i$ for all $i \in I$, $s \succ_i \mu'_i$ implies $s \succ_i \mu_i$. Since $i_k = Top_{s_k}(\{j \in I | s_k \succ_j \mu_j\})$ for each $k$ by definition and $\{j \in I | s \succ_j \mu'_j\} \subset \{j \in I | s \succ_j \mu_j\}$ for any $s$ (and thus in particular for $s_k$), we have $i \not\succ_{s_k} i_k$ for every $i \in \{j \in I | s_k \succ_j \mu'_j\}$. Hence, no one has justified envy to $i_k$ under $\mu'$.

To show that $\mu'$ is balanced, consider a sequence $(r_1, \ldots, r_m)$ such that $s_k \in r_k$ for each $k$. Fix $r \in R$ and let $K(r) = \{k \in \{1, \ldots, m\} | r_k = r\}$. If $K(r) = \emptyset$, then (2.1) is satisfied for $r$ under $\mu'$ because it is satisfied for $r$ under $\mu$. So suppose $K(r) \neq \emptyset$. Let

$$In_r = \{k \in K(r) | r_{k-1} \neq r\} \text{ and } Out_r = \{k \in K(r) | r_{k+1} \neq r\}.$$

Clearly we must have $|In_r| = |Out_r|$. Define $In'_r \subseteq In_r$ and $Out'_r \subseteq Out_r$ by

$$In'_r = \{k \in K(r) | r_{k-1} \neq r \text{ and } i_k \in r\} \text{ and } Out'_r = \{k \in K(r) | r_{k+1} \neq r \text{ and } i_{k+1} \in r\}.$$

On the one hand, the inflow to $r$ has changed from $\mu$ to $\mu'$ by

$$\underbrace{(|In_r| - |In'_r|)}_{\substack{\text{the number of non-}r\text{ students} \\ \text{coming to } r}} \quad - \quad \underbrace{(|Out_r| - |Out'_r|)}_{\substack{\text{the number of non-}r\text{ students} \\ \text{going out of } r}} .$$

On the other hand, the outflow from $r$ has changed from $\mu$ to $\mu'$ by

$$\underbrace{|Out'_r|}_{\substack{\text{the number of } r\text{ students} \\ \text{going out of } r}} \quad - \quad \underbrace{|In'_r|}_{\substack{\text{the number of } r\text{ students} \\ \text{coming to } r}} .$$

Note that these two values are equal because $|In_r| = |Out_r|$. Finally, since (2.1) holds for $r$ under $\mu$, this implies that (2.1) holds for $r$ under $\mu'$.

To show that individual rationality of $\mu$ implies individual rationality of $\mu'$, note first that $\mu'_i \succeq \emptyset$ for each $i \in I$ because $\mu$ is individually rational and $\mu'$ Pareto dominates $\mu$. Moreover, $i \succ_s \emptyset$ for every $s \in S$ and $i \in \mu'_s$ because (i) $i \succ_s \emptyset$ for every $i \in \mu'_s \cap \mu_s$ by the assumption that $\mu$ is individually rational, while (ii) $i \succ_s \emptyset$ for every $i \in \mu'_s \setminus \mu_s$ by the definition of FIG (more specifically, item 1 of Definition 3). $\square$

A.2. **Proof of Theorem 2.**

*Proof.* We prove the contraposition, so fix an iBF $\mu$ and assume that there exists an iBF $\mu'$ that Pareto dominates $\mu$.

**Lemma 1.** *For any* $r \in R$, $|\{i \in I | \mu_i \in r, \mu'_i \notin r\} \cup \{i \in r | \mu_i = \emptyset, \mu'_i \in r' \text{ for some } r' \neq r\}|$ *is equal to* $|\{i \in I | \mu_i \in r' \text{ for some } r' \neq r, \mu'_i \in r\} \cup \{i \in I \setminus r | \mu_i = \emptyset, \mu'_i \in r\}|$.

*Proof.* Since $\mu$ and $\mu'$ are both balanced, the change (from $\mu$ to $\mu'$) of the inflow of students to $r$ and the change of the outflow are the same as each other. The change of the inflow is:

$$\underbrace{|\{i \in I \setminus r | \mu_i \in r' \text{ for some } r' \neq r, \mu'_i \in r\}|}_{\substack{\text{the number of non-}r\text{ students} \\ \text{coming from an outside school to } r}} + \underbrace{|\{i \in I \setminus r | \mu_i = \emptyset, \mu'_i \in r\}|}_{\substack{\text{the number of non-}r\text{ students} \\ \text{coming from being unmatched to } r}} - \underbrace{|\{i \in I \setminus r | \mu_i \in r, \mu'_i \notin r\}|}_{\substack{\text{the number of non-}r\text{ students} \\ \text{going out of } r}}.$$

The change of the outflow is:

$$\underbrace{|\{i \in r | \mu_i \in r, \mu'_i \in r' \text{ for some } r' \neq r\}|}_{\substack{\text{the number of } r \text{ students} \\ \text{going out of } r}} + \underbrace{|\{i \in r | \mu_i = \emptyset, \mu'_i \in r' \text{ for some } r' \neq r\}|}_{\substack{\text{the number of } r \text{ students} \\ \text{who was unmatched but is now matched to a non-}r\text{ school}}}$$
$$- \underbrace{|\{i \in r | \mu_i \in r' \text{ for some } r' \neq r, \mu'_i \notin r' \text{ for any } r' \neq r\}|}_{\substack{\text{the number of } r \text{ students} \\ \text{who was matched to a non-}r\text{ school but now is not}}}.$$

Thus, we have:

(A.1)
$$\underbrace{|\{i \in r | \mu_i \in r, \mu'_i \in r' \text{ for some } r' \neq r\}|}_{\substack{\text{the number of } r \text{ students} \\ \text{going out of } r}} + \underbrace{|\{i \in r | \mu_i = \emptyset, \mu'_i \in r' \text{ for some } r' \neq r\}|}_{\substack{\text{the number of } r \text{ students} \\ \text{who was unmatched but is now matched to a non-}r\text{ school}}}$$
$$+ \underbrace{|\{i \in I \setminus r | \mu_i \in r, \mu'_i \notin r\}|}_{\substack{\text{the number of non-}r\text{ students} \\ \text{going out of } r}} =$$
$$\underbrace{|\{i \in I \setminus r | \mu_i \in r' \text{ for some } r' \neq r, \mu'_i \in r\}|}_{\substack{\text{the number of non-}r\text{ students} \\ \text{coming from an outside school to } r}} + \underbrace{|\{i \in I \setminus r | \mu_i = \emptyset, \mu'_i \in r\}|}_{\substack{\text{the number of non-}r\text{ students} \\ \text{coming from being unmatched to } r}}$$
$$+ \underbrace{|\{i \in r | \mu_i \in r' \text{ for some } r' \neq r, \mu'_i \notin r' \text{ for any } r' \neq r\}|}_{\substack{\text{the number of } r \text{ students} \\ \text{who was matched to a non-}r\text{ school but now is not}}}.$$

Now, recall that $\mu$ is individually rational. Hence, $\mu'_i \succeq_i \mu_i \succeq_i \emptyset$ for every $i$, and thus we have that $\mu_i \in r$ implies $\mu'_i \neq \emptyset$. Therefore,

$$\{i \in r | \mu_i \in r, \mu'_i \in r' \text{ for some } r' \neq r\}| = |\{i \in r | \mu_i \in r, \mu'_i \notin r\}$$

Also, for the same reason, we have

$$\{i \in r | \mu_i \in r' \text{ for some } r' \neq r, \mu'_i \notin r' \text{ for any } r' \neq r\} = \{i \in r | \mu_i \in r' \text{ for some } r' \neq r, \mu'_i \in r\}.$$

Hence, (A.1) is equivalent to

$$|\{i \in r|\mu_i \in r, \mu_i' \notin r\}|+|\{i \in r|\mu_i = \emptyset, \mu_i' \in r' \text{ for some } r' \neq r\}|+|\{i \in I\backslash r|\mu_i \in r, \mu_i' \notin r\}| =$$
$$|\{i \in I \setminus r|\mu_i \in r' \text{ for some } r' \neq r, \mu_i' \in r\}| + |\{i \in I \setminus r|\mu_i = \emptyset, \mu_i' \in r\}|$$
$$+|\{i \in r|\mu_i \in r' \text{ for some } r' \neq r, \mu_i' \in r\}|,$$

or

$$|\{i \in I|\mu_i \in r, \mu_i' \notin r\}| + |\{i \in r|\mu_i = \emptyset, \mu_i' \in r' \text{ for some } r' \neq r\}| =$$
$$|\{i \in I|\mu_i \in r' \text{ for some } r' \neq r, \mu_i' \in r\}| + |\{i \in I \setminus r|\mu_i = \emptyset, \mu_i' \in r\}|.$$

Since the two terms in each side of the above equation are disjoint from each other, we have

$$|\{i \in I|\mu_i \in r, \mu_i' \notin r\} \cup \{i \in r|\mu_i = \emptyset, \mu_i' \in r' \text{ for some } r' \neq r\}| =$$
$$|\{i \in I|\mu_i \in r' \text{ for some } r' \neq r, \mu_i' \in r\} \cup \{i \in I \setminus r|\mu_i = \emptyset, \mu_i' \in r\}|.$$

This completes the proof. □

Consider the following graph, in which the only agents associated with arrows are the schools in $r$ and students who are matched to a school in $r$ under $\mu$ and students living in $r$ who are unmatched under $\mu$. First, each $s \in r$ points to each student $i \in \mu_s \setminus \mu_s'$. Then, for each student who was pointed to by some school in $r$ and each student living in $r$ who are unmatched under $\mu$, let her point to the school $\mu_i'$ if $\mu_i' \in r$. Moreover, by Lemma 1, there is a one-to-one and onto mapping from $\{i \in I|\mu_i \in r, \mu_i' \notin r\} \cup \{i \in r|\mu_i = \emptyset, \mu_i' \in r' \text{ for some } r' \neq r\}$ to $\{i \in I|\mu_i \in r' \text{ for some } r' \neq r, \mu_i' \in r\} \cup \{i \in I \setminus r|\mu_i = \emptyset, \mu_i' \in r\}$. Take one such mapping $\phi$. Then, for each $i \in \{i \in I|\mu_i \in r, \mu_i' \notin r\} \cup \{i \in r|\mu_i = \emptyset, \mu_i' \in r' \text{ for some } r' \neq r\}$, let $i$ point to $\mu_{\phi(i)}'$. This defines a directed graph, denoted $\mathcal{G}(\mu, \mu', r)$, in which only schools in $r$, students matched to a school in $r$ under $\mu$, and the students of $r$ who are unmatched under $\mu$ may be associated with arrows.

Let $I' := \{i \in I|\mu_i' \succ_i \mu_i\}$. By the assumption that $\mu'$ Pareto dominates $\mu$, we have $I' \neq \emptyset$. For any matching $\tilde{\mu}$, let $S(\tilde{\mu}) := \{s \in S|s = \tilde{\mu}_i \text{ for some } i \in I'\}$.

**Lemma 2.** *Suppose $s \in S(\mu) \setminus S(\mu')$. Then there exists a school $s' \in r(s) \cap S(\mu')$ such that $|\mu_{s'}| < q_{s'}$.*

*Proof.* Take an arbitrary $s$ such that $s \in S(\mu) \setminus S(\mu')$ (If there is no such school, we are done). Starting from this school, follow the arrows in $\mathcal{G}(\mu, \mu', r(s))$ in an arbitrary manner without passing the same student twice (note that there is an outgoing arrow from $s$). Since there are a finite number of students, there is $s'$ such that there is no more outgoing arrow from $s'$ to a student who has not appeared in the path (note that, by definition, this path cannot end at any student). This implies that the number of students who are in $\mu'_{s'} \setminus \mu_{s'}$ is larger than the number of students who are in $\mu_{s'} \setminus \mu'_{s'}$ by at least one. Hence, we have $q_{s'} \geq |\mu'_{s'}| > |\mu_{s'}|$. Since $s' \in r(s)$ and $s' \in S(\mu')$ by the definition of the graph, this completes the proof. □

**Lemma 3.** *Suppose $\mu_i = \emptyset$ and $i \in I'$. Then there exists a school $s' \in r(i) \cap S(\mu')$ such that $|\mu_{s'}| < q_{s'}$.*

*Proof.* Suppose there is $i$ such that $\mu_i = \emptyset$ and $i \in I'$. Starting from this student $i$, follow the arrows in $\mathcal{G}(\mu, \mu', r)$ in an arbitrary manner without passing the same student twice (note that there is an outgoing arrow from $i$). Since there are a finite number of students, there is $s'$ such that there is no more outgoing arrow from $s'$ to a student who has not appeared in the path (note that, by definition, this path cannot end at any student). This implies that the number of students who are in $\mu'_{s'} \setminus \mu_{s'}$ is larger than the number of students who are in $\mu_{s'} \setminus \mu'_{s'}$ by at least one. Hence, we have $q_{s'} \geq |\mu'_{s'}| > |\mu_{s'}|$. Since $s' \in r(i)$ and $s' \in S(\mu')$ by the definition of the graph $\mathcal{G}(\mu, \mu', r(i))$, this completes the proof. □

Next, we define a graph, denoted $\mathcal{G}(\mu, \mu')$, as follows. In this graph, only students in $I'$ and schools in $S(\mu')$ may be associated with arrows. Formally, for any $s \in S(\mu')$, consider the set of students in $I'$ who strictly prefer $s$ to their match at $\mu$, i.e., $D^{\mu}_{s}(I') := \{i \in I' | s \succ_i \mu_i\}$, and let $Top_s(D^{\mu}_{s}(I'))$ point to $s$. Note that $D^{\mu}_{s}(I')$ is nonempty by the definitions of $I'$, $\mu$ and $\mu'$, and thus for any $s \in S(\mu')$, there exists some $i \in I'$ who points to $s$. Next, consider any $i \in I'$.

(1) If $\mu_i \in S(\mu')$, then let $\mu_i$ point to $i$.
(2) If $\mu_i \in S \setminus S(\mu')$, then $\mu_i \in S(\mu) \setminus S(\mu')$ by the definition of $S(\mu)$ and hence by Lemma 2, there exists a school $s' \in r(\mu_i) \cap S(\mu')$ such that $|\mu_{s'}| < q_{s'}$. Let any such school $s'$ point to $i$.
(3) If $\mu_i = \emptyset$, then by Lemma 3, there exists $s \in r(i) \cap S(\mu')$ such that $|\mu_s| < q_s$. Let any such school $s$ point to $i$.

The graph must have a cycle because each school is pointed to by a single student, and each student is pointed to by at least one school.[19] Pick an arbitrary cycle and call it $\mathcal{F}^*$.

**Lemma 4.** *$\mathcal{F}^*$ is a FIG cycle.*

[19] To find a cycle, take an arbitrary school and find the student pointing to that school. Then find the school pointing to that student. Then find the student pointing to that school, etc., until we find a school or a student that has already been visited. Since there are only finitely many students, this procedure ends in finite steps.

*Proof.* Let $\mathcal{F}^* = (i_1, s_1, \ldots, i_m, s_m)$.

It is straightforward to check that the last two conditions of a cycle are satisfied: By construction, each student appears only once in $\mathcal{F}^*$. Given this, since the pointing rule for $\mathcal{G}(\mu, \mu')$ implies that each school is pointed to only by a single student, each school appears at most once in $\mathcal{F}^*$.

To complete the proof, it suffices to show that $\mathcal{G}(\mu, \mu')$ is a subset of the FIG on $\mu$. To show this, first we establish that, for any $k \in \{1, \ldots, m\}$, $i_k$ points to $s_k$ according to the definition of pointing used for FIG. To do so, it suffices to show that $i_k = Top_{s_k}(D^{\mu}_{s_k})$ and $i_k \succ_{s_k} \emptyset$ (Case 1 of Definition 3). The latter holds for the following reason: we have $i \in D^{\mu}_{s_k}(I')$ for some $i \in \mu'_{s_k}$ because of the definitions of $S(\mu')$ and $I'$ and the fact that $s_k \in S(\mu')$. Hence, by individual rationality of $\mu'$, it follows that $i_k = Top_{s_k}(D^{\mu}_{s_k}(I')) \succeq_{s_k} i \succ_{s_k} \emptyset$. To show the former, note first that, by construction, $i_k = Top_{s_k}(D^{\mu}_{s_k}(I'))$ and hence $i_k \succ_{s_k} i$ for any $i \in (D^{\mu}_{s} \cap I') \setminus \{i_k\}$. Next, consider any $i \in D^{\mu}_{s_k} \setminus I'$. Because $\mu'_{i'} = \mu_{i'}$ for any $i' \in I \setminus I'$ by the definition of $I'$, it follows that $i \in D^{\mu'}_{s_k}$. This and the assumption that $\mu'$ is fair imply $j \succ_{s_k} i$ for every $j \in \mu'_{s_k}$. By the construction of the cycle, $i_k \succeq_{s_k} j$ for every $j \in \mu'_{s_k} \setminus \mu_{s_k} \neq \emptyset$ (the nonemptiness holds because $s_k \in S(\mu')$). Thus, we have $i_k \succ_{s_k} i$. Therefore, we have $i_k \succ_{s_k} i$ for any $i \in D^{\mu}_{s} \setminus \{i_k\}$, which implies $i_k = Top_{s_k}(D^{\mu}_{s_k})$.

Second, we consider three cases of the definition of $\mathcal{G}(\mu, \mu')$ to show that, for any $k \in \{1, \ldots, m\}$, $s_k$ points to $i_{k+1}$ according to the definition of pointing used for FIG (with $m + 1 = 1$). Suppose first that $i_{k+1}$ and $s_k$ satisfy the condition described in Case 1 of the definition of $\mathcal{G}(\mu, \mu')$. This implies that $i_{k+1}$ and $s_k$ satisfy the assumption in Case 2a of Definition 3. Next, consider Case 2 of the definition of $\mathcal{G}(\mu, \mu')$. In this case, $|\mu_{s_k}| < q_{s_k}$ and $\mu_{i_{k+1}} \in r(s_k)$ hold, which satisfies the condition in Case 2b of Definition 3. Finally, consider Case 3 of the definition of $\mathcal{G}(\mu, \mu')$. In this case, $|\mu_{s_k}| < q_{s_k}$, $i_{k+1} \in r(s_k)$, and $\mu_{i_{k+1}} = \emptyset$ hold, which again satisfies the condition in to Case 2b of Definition 3. This completes the proof. □

Lemma 4 completes the proof. □