Equal opportunities in school choice settings

Domenico Moramarco* Umutcan Salman †

April 24, 2023

Abstract

We investigate the possibility of integrating a notion of fairness, inspired by the equality of opportunity literature, into the centralized school choice setting. By doing so, we enlarge the list of available tools to implement equality of opportunity in education. We enrich the standard school choice setting with the notion of educational outcome or quality of a matching between a school and a student. In our framework, fairness considerations are made by a social evaluator based on the match quality distribution. We investigate the compatibility between our notion of fairness, a notion of efficiency based on aggregate match quality, and the standard notion of stability. To overcome some of the identified incompatibilities, we propose two alternative approaches. The first one is a linear programming solution to maximize fairness under stability constraints. The second approach weakens fairness and efficiency to define a class of opportunity egalitarian social welfare functions that evaluate stable matchings. We then describe an algorithm to find the stable matching that maximizes social welfare.

1 Introduction

Equality of opportunity is a fundamental principle of a just society, and education is a key driver of individual opportunity. As such, the economic literature has often suggested investments in public education as a policy to fight inequality of opportunity. In recent years, centralized school choice mechanisms have become increasingly popular as a way to efficiently allocate students to schools. While concerns for minority groups have been addressed in the literature on school choice, there remains a gap in understanding how these mechanisms can be used to achieve equality of educational opportunity. This paper aims to fill this gap by investigating the potential of centralized school choice mechanisms to promote equality of educational opportunity.

Equality of opportunity. Many political philosophers (Rawls, 1971; Sen *et al.*, 1980; Dworkin, 1981a,b) have debated whether all inequalities should be considered unacceptable or if there are certain inequalities that a fair society should tolerate and preserve.

*University of Bari - Department of Economics and Finance. Email: domenico.moramarco@uniba.it

†Université libre de Bruxelles - European Center of Advanced Research in Economics and Statistics. Email: umutcan.salman@ulb.be

Economists (Cappelen *et al.*, 2007, 2013; Alesina *et al.*, 2017) have also investigated preferences for redistribution and tried to identify inequalities that individuals consider unfair. Additionally, studies (Marrero & Rodríguez, 2013) have distinguished between inequalities that are good or bad for economic development.

All these scholars converge on the idea that inequalities stemming from individual characteristics that are out of control or responsibility are unfair and detrimental to socio-economic development. The term "equality of opportunity" (EOp) is used in the literature to refer to both the absence of such inequalities and the field of study that developed after the contributions of Roemer (1993), Bossert (1995) and Fleurbaey (1994). In a nutshell, EOp means levelling the playing field so that the final outcome of each individual is ultimately due to their own choices. As also argued by Roemer & Ünveren (2017), a key source of observed inequality of opportunity (IOp) in income is the education premium that some students gain from attending better schools. Researchers such as Corak (2013) have emphasized the importance of primary and secondary education in shaping opportunity and social mobility. Thus, it is not surprising that the key policy recommendation for equalizing opportunities and improving social mobility is to increase investments in public education. However, investing in public education is costly and governments have limited budgets.

Centralized school choice. In many countries (e.g. Sweden, Chile, France, Turkey and the Netherlands) and cities (e.g. New Orleans, New York, and Boston) around the globe, students are assigned to public schools through centralized school choice systems. A school choice system is a two-sided matching market in which there are students (or their parents) on one side of the market, and public schools on the other side of the market. In this system, students (or their parents) submit their preferences for public schools, and the schools rank students based on certain criteria. The criteria used to rank students can vary, but they often include factors such as distance from the school, socioeconomic status, and academic performance.

With no (or fixed) tuition fees for public schools, the central authority uses an algorithm to match students with schools while respecting preferences and priorities. The goal of the educational authority is to design an algorithm that finds assignments with desirable properties such as stability and Pareto efficiency. Stability is the central fairness notion in two-sided matching markets; it requires that in the final assignment there exist no pair of student-school that would prefer each other to their current matching. Pareto efficiency can be regarded as a welfare concept in matching markets; an assignment is said to be Pareto efficient if there is no other assignment that makes a student better off without hurting some other student. The incompatibility between these properties is well-established in the literature (see Abdulkadiroğlu & Sönmez, 2003). A well known (compromise) solution to this impossibility is the Deferred Acceptance (DA) algorithm (Gale & Shapley, 1962), that provides the most efficient assignment for students among all stable ones. This algorithm, which plays a central role in what follows, is also used in some school choice settings; for example, after Abdulkadiroğlu & Sönmez

(2003), the educational authorities in New York City are adopting DA to assign students to public high schools.

Educational opportunity. In school choice settings, affirmative action policies are often implemented to promote disadvantaged students or reduce segregation in schools. This is referred to as controlled school choice. There are three main types of affirmative action policies in the literature: majority quotas (Abdulkadiroğlu & Sönmez, 2003), minority reserves (Hafalir *et al.*, 2013), and priority-based affirmative action. Majority quotas limit the number of majority students that can be admitted to a school, minority reserves reserve seats for minority students, and priority-based affirmative action prioritizes minority students in the admissions process. While it has been shown (see Kojima, 2012; Doğan, 2016; Afacan & Salman, 2016) that these policies may have negative effects on the targeted groups, focusing on few particular minorities may lead us to neglect other disadvantaged groups. This paper aims to implement a more holistic approach, typical of the EOp paradigm, which takes into account multiple groups simultaneously.

Abdulkadiroğlu *et al.* (2020) find that parent's preferences do not always align with the quality of education their children can receive from given schools. Therefore, in this paper, we follow Abdulkadiroglu *et al.* (2021) in considering the quality of a student-school match as the relevant outcome for the central authority to evaluate assignments. More precisely, we consider match quality to be the potential educational outcome a student can receive by attending a school. Although we will discuss particular cases, match quality is assumed to be measured by a central authority for any possible student-school pair. Under such assumption, equalizing educational opportunities for students coincides with equalizing opportunities for match quality. This paper focuses on the prominent interpretations of the EOp paradigm (see Ramos & Van de Gaer, 2016, for a survey), which defines social justice as equality of expected outcomes across groups of individuals with similar circumstances (such as, for example, gender, parents' education, and ethnicity) that are out of their control.

Methodology. In a standard school choice setting, augmented with the a partition of students into types and a measure of match quality, we discuss three desirable properties for a school assignment: stability, efficiency and fairness. While the former corresponds with the standard requirement in matching theory, the second property has already been proposed by Abdulkadiroglu *et al.* (2021). The fairness requirement is, however, new and is based on the idea of minimizing the inequality in types' expected educational outcome. We highlight the incompatibility between these three requirements and provide two approaches to solve it. The first solution consists in using linear programming to identify the fair, or the efficient, allocation within the set of stable ones. The second approach is a normative one which defines a family of opportunity egalitarian social welfare functions, in line with Peragine (2004), and an algorithm - the Stable Opportunity Egalitarian (SOE) - based on Cheng *et al.* (2008), which maximize social welfare within the set of stable matchings.

Contribution. We contribute to the literature in three ways. First, to the best of our knowledge, this is the first paper that introduces the concept of Equality of Opportunity à la Roemer (1998) in a centralized school choice setting. Second, despite the well-known and documented importance of public education in enhancing opportunities, the EOp literature has hardly included centralized school admission systems within the set of policy recommendations. This paper shows the possibility of designing school allocation mechanism inspired to opportunity egalitarian fairness principles. Finally, our use of Cheng *et al.* (2008)'s proposal in the SOE algorithm shows how this procedure can be implemented in many other economic settings to maximize utilitarian social welfare functions over the set of stable matchings.

Example. We conclude this section with a simple example of how affirmative action policies may not align with opportunity egalitarianism. Consider a simple scenario with six students and three schools, each offering two seats. The schools differ in quality, which we approximate by using the average probability of a student receiving an offer from a top college. School 1 has probability of 4%, school 2 has 10%, and school 3 has 20%. Suppose that students 1 and 2 belong to the most disadvantaged population subgroup, students 5 and 6 belong to the most advantaged subgroup, while students 3 and 4 belong to population a subgroup within these two extremes. Let preferences of students (i) and priority rankings of schools (s) be:

$$\begin{array}{lcl} i_1 = i_2 & : & s_3, s_1, s_2 \\ i_3 = i_4 = i_5 = i_6 & : & s_3, s_2, s_1 \\ s_1 & : & i_6, i_5, i_1, i_2, i_3, i_4 \\ s_2 = s_3 & : & i_6, i_5, i_4, i_3, i_2, i_1 \end{array}$$

where, for example, s_3 is the most preferred school for i_1 and i_2, and i_4 is the least preferred student for s_1. Following a standard approach in the literature, we represent the opportunity distribution with a vector $\mathbf{p} = (p_1, p_2, p_3)$, where p_i is the expected probability for a random student of group i of receiving an offer from a top college. Let us consider the opportunity distributions generated by the following three allocations:

$$\begin{bmatrix} & DA & \\ s_1 & \leftarrow & i_1, i_2 \\ s_2 & \leftarrow & i_3, i_4 \\ s_3 & \leftarrow & i_5, i_6 \end{bmatrix} \begin{bmatrix} & EOp & \\ s_1 & \leftarrow & i_1, i_2 \\ s_2 & \leftarrow & i_3, i_4 \\ s_3 & \leftarrow & i_5, i_6 \end{bmatrix} \begin{bmatrix} & AA & \\ s_1 & \leftarrow & i_1, i_2 \\ s_2 & \leftarrow & i_3, i_4 \\ s_3 & \leftarrow & i_5, i_6 \end{bmatrix}$$

where DA is the standard Gale & Shapley (1962)'s Deferred Acceptance (DA) algorithm explained in Section 4, EOp is obtained by exchanging the school assignment of students 2 and 6, and AA is obtained when school 3 reserves one seat to students from the most disadvantaged group (a minority quota) and DA is implemented. Observe that $\mathbf{p}^{DA} = (4, 10, 20)$, $\mathbf{p}^{EOp} = (12, 10, 12)$ and $\mathbf{p}^{AA} = (12, 7, 15)$. The Gini coefficient of the opportunity distribution is 0.31 for $\mathbf{p}^{DA}$, 0.16 for $\mathbf{p}^{AA}$ and 0.04 for $\mathbf{p}^{EOp}$. The comparison in terms of equality of opportunity is neat, revealing that Deferred Acceptance does not guarantee distributional fairness. It is interesting to also notice how, because of the focus on a single group, the affirmative action policy, in this example, puts the burden of fairness also on the second group who is not as advantaged as the third one. This consequence may be normatively unappealing.[1] By giving importance to all groups, the opportunity egalitarian approach we implement in this paper tries to also limit this last issue.

Outline. The paper is organized as follows. Section 2 introduces the basic notation and definitions. Section 3 introduces the three desirable properties, discusses their compatibility and the complexity of optimizing over the set of stable matchings. Section 4 introduces a family of opportunity egalitarian social welfare functions and the SOE algorithm. Section 5 discusses the positive and normative implication of the algorithm. Section 6 concludes.

2 The school choice setting

In this section we describe the theoretical setting and introduce some preliminary definitions. An instance of centralized school choice problem, with EOp components, is a tuple $\mathcal{I} = \langle I, S, P, \succ, q, T, \trianglelefteq, U \rangle$, where $I = \{i_1, i_2, ..., i_{|I|}\}$ is the set of students and $S = \{s_1, s_2, ..., s_{|S|}\}$ is the set of schools. We denote $P= (P_{i_1}, ..., P_{i_{|I|}})$ the students' preferences profile such that, for all $i \in I$ and for all $s, s' \in S$, $s \ P_i \ s'$ means that i strictly prefers s to s'. We assume preference profiles to be complete and strictly linear. The school's priority profile is $\succ= (\succ_{s_1}, ..., \succ_{s_{|S|}})$; for all $s \in S$, $\succ_s$ is the complete and strictly linear priority ranking of school $s \in S$ over I, so that $i \succ_s i'$ means that i has higher priority than i' of being admitted to s. The vector $q = (q_{s_1}, ..., q_{s_{|S|}})$ is the quota profile of schools, so that each school $s \in S$ can admit at most $q_s \in \mathbb{N}_{++}$ students.

We assume that the population of students can be partitioned in mutually exclusive subgroups which, following the EOp literature,(Roemer, 1998) we call types. We denote $T = \{t_1, t_2, ..., t_{|T|}\}$ the set of types (or type partition). As in many applications (Atkinson & Bourguignon, 1982; Peragine, 2002, 2004), we assume the existence of a complete and transitive pre-order of types such that $t \trianglelefteq t'$ means that students of type $t \in T$ are not more advantaged (or have weakly higher needs) than those of $t' \in T$.

The last key ingredient is the educational outcome or *match quality*. We assume that there exists a match quality matrix $U \in \mathbb{R}^{|I| \times |S|}$ such that each cell (i, s) of U represents the potential[2] educational outcome of student i from attending school s. We assume match quality to be comparable across students and schools. For convenience, abusing notation, will sometimes refer to a match quality function $U : I \times S \to \mathbb{R}$, whose value $U(i, s)$ is the match quality of the student-school pair (i, s), which coincides with the relative entry of the match quality matrix. We assume U to be exogenously given by a central authority or an evaluator with sufficient information to assess potential

[1] See also Students for Fair Admissions, Inc. v. President and Fellows of Harvard College (Docket 20–1199), for a recent real life example.

[2] The outcome is potential because we do not assume to observe student's future effort once enrolled.

educational outcomes.

Our assumption does not exclude two particular ways of measuring match quality, which find support both in the school choice and in the EOp literature. The first approach consists in measuring potential educational outcomes according to individual preferences, so that the better preferred school provides higher match quality than a less preferred one. Such a measure can rely on the assumption that attending the favorite school boosts motivation and potential outcomes of students, as well as on the idea that students (or parents) have sufficient information to assess schools' quality. As we will discuss afterward, from a normative perspective, such a match quality measure is in line with the opportunity egalitarian principle of holding individual responsible for their preferences. A second approach to assess match quality relies on schools' priority rankings; these are often based on previous educational outcomes or test scores that can be indicative of the potential educational outcome. At the same time, schools may have a better understanding, based on past experience, of the way students with different abilities respond to their particular teaching methods. These two criteria for assessing match quality have, however, their drawbacks: students (or parents) may not have full knowledge of the school characteristics; students' preferences may be influenced by external factors that are not relevant (or can be detrimental) for educational outcome[3]; schools may have incentives to admit students with better parental background or prefer a certain student composition in order to preserve a status.

An empirical way of defining match quality may consist in taking the average income, or higher education achievement, of other students that attended a given school. For example, one may have information on the income of workers, form a particular ethnic group, that attended school $s \in S$ in the previous years and use it to define the match quality of future students of s belonging to the same ethnic group.

Another criterion to define match quality may rely on rankings by independent authorities or organization. For example, a quick search on the net can provide future university students with the ranking of the best universities and departments in the world. Such rankings are likely to exist for smaller geographical areas and other education degrees, so that the educational authority can assign match quality 1 to the worst school in the area, 2 to the second worst, 3 to the next one and so on. This is also in line with the idea that two students attending the same school would get the same educational outcome. For future reference, we refer to the following definition.

Definition 1. The function $U : I \times S \to \mathbb{R}$ is a match quality measure such that $U(i, s)$ is the potential educational outcome of student i from attending school s.

A preference-based match quality measure is a function $U_P : I \times S \to \mathbb{R}$ such that, for all $i, i' \in I$ and $s, s' \in S$, sP_is' implies $U_P(i, s) > U_P(i, s')$, and $P_i(s) = P_{i'}(s')$ implies $U_P(i, s) = U_P(i', s')$.

A priority-based match quality measure is a function $U_\succ : I \times S \to \mathbb{R}$ such that, for all $i, i' \in I$ and $s, s' \in S$, $i \succ_s i'$ implies $U_\succ(i, s) > U_\succ(i', s)$, and $\succ_s(i) = \succ_{s'}(i')$

[3]For example, a student may prefer the closest school even if it has low quality, rather than a better school at five minutes walking distance.

implies $U_{\succ}(i,s) = U_{\succ}(i',s')$.

A school-based match quality measure is a function $U_S : I \times S \to \mathbb{R}$ such that, for all $i, i' \in I$ and $s \in S$, $U_S(i,s) = U_S(i',s)$.

A centralized school choice setting is a two-sided matching market where there is no price mechanism to clear the market. Hence, an educational authority must apply a matching algorithm to assign students to the available slots in public schools. The result of such an algorithm is an *assignment*,[4] which we express as a function μ from the set $I \cup S$ into the power set of $I \cup S$ such that: *(i)* for all $i \in I$, $\mu(i) \in S$; *(ii)* for all $s \in S$, $|\mu(s)| \leq q_s$ and $\mu(s) \subseteq I$; *(iii)* $\mu(i) = s$ if and only if $i \in \mu(s)$. In words, $\mu(i) = s$ means that student i is enrolled to school s, and $\mu(s)$ denotes the subset of students admitted to school s. A student i is assigned to a school s if and only if student i is one of the students that school s admits. We assume that, in any final assignment, each student is assigned to a school and all the students are entitled to a seat in any schools.

As in Cheng *et al.* (2008), we assume match quality to satisfy the following *independence property.*

Assumption 1 (IND)**.** For all $i \in I$ and $s \in S$, the match quality $U(i,s)$ is function of i and s alone.

A key implication of IND is that $U(i, \mu(i))$ does not depend on the assignment of any other student $j \in I/\{i\}$. This makes it difficult to account for peer effects when measuring match quality. It is worth underlining though that one can still define a match quality measure that accounts for all the peer effects stemming from interactions with other cohorts' students already enrolled, without violating IND. To put it differently, IND requires match quality to depend only on the characteristics of student and school *prior* to the matching. Therefore, the quality of future matches can be influenced by the characteristics of the current student body.

Similar reasoning concerns the issue of school segregation. As it will be clearer later, depending on the match quality matrix, school segregation may be compatible with equality of opportunity. This is, however, only in part due to independence. While IND prevents us from adapting match quality to the demographics of students admitted simultaneously, the ethnic composition of the previously admitted students can be taken into account by the central authority that assesses match quality.

3 Desirable properties

This section defines three desirable properties of a school assignment, discusses their incompatibility and proposes an intermediary solution based on linear programming methods.

It is standard in the matching literature to require stability of the final allocation. A stable matching μ, in a two sided matching market, can be formalized as follows.

[4]Throughout the text we will use the terms assignment, matching and allocation as synonymous.

Axiom 1 (Stability)**.** There is no student-school pair $(i, s) \in I \times S$ such that $s \ P_i \ \mu(i)$ and either $|\mu(s)| < q_s$ or $|\mu(s)| = q_s$ and $i \succ_s i'$ for some $i' \in \mu(s)$.

In words, an assignment μ is stable if there is no student-school (i, s) pair such that student i prefers school s to his assignment and either school s has an empty slot or school s has no available slot and it ranks i higher than some of its admitted student.

Stability guarantees the respect for both schools' priorities and student's preferences. In a non-stable assignment there exist a student i that prefers school s to her current assignment and has higher priority than a student j admitted at s. This situation, which would be deemed unfair in the social choice literature, gives student i the right to object the final assignment and complaint to the public authority. This type of complaints can be a burden for the authorities which may find stability to be desirable also from a practical point of view. In line with the prominent approaches in the literature, we keep stability as the necessary requirement for any desirable school allocation.

As also suggested by Abdulkadiroglu *et al.* (2021), a second desirable property for an assignment μ is the maximization of the aggregate match quality. This is a natural efficiency requirement that is supported both within and beyond the fairness literature. If U is the potential educational outcome, most utilitarian social planners would aim to maximize the (potentially transformed) sum of educational outcomes, as this increases human capital accumulation, growth, and social welfare.

Axiom 2 (Efficiency)**.** There is no other assignment μ' such that

$$\sum_{i \in I} U\left(i, \mu'(i)\right) > \sum_{i \in I} U\left(i, \mu(i)\right).$$

The next desirable property is a fairness requirement inspired to the EOp paradigm. For any match quality measure U and assignment μ, let

$$\bar{u}\left(t, \mu\right) = \frac{1}{|t|} \sum_{i \in t} U\left(i, \mu\left(i\right)\right)$$

be the educational opportunity of an individual belonging to type $t \in T$. This definition of educational opportunity follows the ex-ante approach to EOp (Fleurbaey & Peragine, 2013; Roemer & Trannoy, 2015; Ramos & Van de Gaer, 2016), which evaluates individual opportunities in terms of expected outcomes conditional on individual characteristics. Hence, equalizing opportunities coincides with minimizing the differences between expected match quality across types. The third desirable property naturally follows from this.

Axiom 3 (Fairness)**.** There is no other assignment μ' such that

$$\sum_{j,k=1}^{|T|} \left|\bar{u}\left(t_j, \mu'\right) - \bar{u}\left(t_k, \mu'\right)\right| < \sum_{j,k=1}^{|T|} \left|\bar{u}\left(t_j, \mu\right) - \bar{u}\left(t_k, \mu\right)\right|.$$

It comes with no surprise that, for any school choice instance $\mathcal{I}$, there always exist an efficient, a fair and a stable allocation. The existence for the first two follows from the fact that the set of possible assignments induced by an instance $\mathcal{I}$ is finite, so that there always exists an assignment that maximizes total match quality or minimizes the differences across types. The second claim has been shown by Gale & Shapley (1962) by designing DA, which always provides a stable matching.

Observe that Stability is a requirement imposed on how the assignment relates with preferences and priority rankings. The other two requirements look at how the assignment performs in terms of match quality, which may not represent preferences or priority rankings. It is easy to see how one can find instances to show the incompatibility between Stability and any of the other two requirements. As it is often the case in welfare economics, incompatibility exists also between Efficiency and Fairness. To see this clearly, consider the following example.

Example. Consider the following school choice instance: where $I = \{i_1, i_2, i_3, i_4\}$, $S = \{s_1, s_2\}$, $T = \{t_1, t_2\}$, $t_1 = \{i_1, i_2\}$, $t_2 = \{i_3, i_4\}$, $q = (q_{s_1}, q_{s_2}) = (2, 2)$,

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	$\succ_{s_1}$	$\succ_{s_2}$
s_1	s_1	s_1	s_1	i_1	i_1
s_2	s_2	s_2	s_2	i_2	i_2
				i_3	i_3
				i_4	i_4

U	s_1	s_2
i_1	1	3
i_2	2	3
i_3	3	1
i_4	2	1

There exist one stable, $\mu_S = ((i_1, s_1), (i_2, s_1), (i_3, s_2), (i_4, s_2))$, one efficient, $\mu_E = ((i_1, s_2), (i_2, s_2), (i_3, s_1), (i_4, s_1))$, and one fair, $\mu_F = ((i_1, s_1), (i_2, s_2), (i_3, s_1), (i_4, s_2))$ matchings. Any two of these matchings do not coincide.

The following Lemma defines sufficient conditions for the existence of a matching that is both efficient and fair.

Lemma 1. Let $\mathcal{I}$ be such that: *(i)* $|t| = |t'|$ for any $t, t' \in T$, and *(ii) there exists* $z_s \in \mathbb{N}_{++}$ *such that* $q_s = z_s|T|$ *for all* $s \in S$. If $U : I \times S \to \mathbb{R}$ is a school-based match quality measure, then there always exist an assignment μ that is both fair and efficient.

Proof. The proof consists in providing a simple procedure to obtain an assignment with no difference in type's expected match quality. Without loss of generality, let schools be ordered such that, for all $j \in \{1, ..., |S| - 1\}$, $U_S(\cdot, s_j) \geq U_S(\cdot, s_{j+1})$. For all $s \in S$, let $z_s^* = \max\{z \in \mathbb{N}_{++} : q_s = z|T|\}$. The procedure can be described as follows:

- *Step 1*: Assign $z_{s_1}^*$ students from each type to s_1. If there are unassigned students, go to following step, otherwise stop.
- *Step k:* Assign $z_{s_k}^*$ students from each type to s_k. If there are unassigned students, go to following step, otherwise stop.

Let μ denote the resulting allocation. To see that μ satisfies fairness, let us denote θ the match quality of an unassigned student, then at each step k we have

$$\bar{u}(t,\mu) = \frac{1}{|t|}\left(\sum_{j=1}^{k} z^*_{s_j} U_S(\cdot, s_j) + \left(|t| - \sum_{j=1}^{k} z^*_{s_j}\right)\theta\right)$$

for all $t \in T$. At the final step, k will be equal to $|S|$. Hence $\left(\left(|t| - \sum_{j=1}^{k} z^*_{s_j}\right)\theta\right)$ will be zero and for each type the average match quality will be the same.

The second part of the statement follows from the fact that $q_s = z_s|T|$ implies $z^*_s = z_s$ so that at each step k, none of the s_{k-1} schools has free seats. Consequently, there is no margin for improving the total match quality.

□

The previous lemma tells us that in some school choice instances it is possible to find an assignment that is both fair and efficient. Observe that, if we have schools with the same U_S, then condition (ii) in the previous lemma can be weakened to impose the *sum* of their quotas to be proportional to $|T|$. Notice also that, if all schools have the same quality, then fairness is trivially compatible with complete school segregation. Clearly, if there are schools with low quality, and the available seats in the better schools are sufficient to assign all students in a fairness coherent way, then (ii) must hold only for a subset of S.

The conditions under which one obtains the previous result are however strong. Moreover, there is no guarantee that the fair assignment will also be stable. In the rest of the paper we will make stability the necessary requirement an assignment must satisfy in order to be considered among the possible final allocations. In other words, we will be looking for either fair or efficient assignments within the set of stable ones. This is not a trivial problem because constructing the set of all stable assignments can be difficult for some school choice instances. More clearly, the number of stable assignments for some instances can be very big. Hence finding all stable assignments is not a computationally tractable problem.[5]

Despite the computational complexity, finding the fair (resp. efficient) assignment among the stable matchings can be written as an optimization problem, solvable with standard linear programming procedures.

Formally, for any matching μ, let $x \in \{0,1\}^{|I|\times|S|}$ be such that $x_{(i,s)} = 1$ if $\mu(i) = s$ and $x_{(i,s)} = 0$, The linear programming (LP) below, finds a fair assignment among the stable ones.

[5] See Gusfield & Irving (1989).

$$
\begin{array}{lll}
\min & \sum_{j,k=1}^{|T|} \left| \sum_{i \in t_j, s \in S} \frac{1}{|t_j|} U(i,s)\, x_{(i,s)} - \sum_{i \in t_k, s \in S} \frac{1}{|t_k|} U(i,s)\, x_{(i,s)} \right| & \\
\text{s. t.} & \sum_{s \in S} x_{(i,s)} = 1 & \forall i \in I \\
& \sum_{i \in I} x_{(i,s)} \leq q_s & \forall s \in S \\
& \sum_{(i',s) \in \mathcal{C}(i,s)} x_{(i',s)} \geq q_s & \forall s \in S, \; \forall \mathcal{C}(i,s) \in \mathbb{C}_s
\end{array}
$$

Specifically, the first constraint imposes that any student is assigned to one school, the second checks that the assignment respects the quota of each school. The last condition of LP is checking if, for each student-school pair, (i, s) either student i is matched to a better school than s or the number of students who are ranked better than i in $\succ_s$ and admitted by the school s is equal to $|q_s|$. Hence, this condition, provided by Baïou & Balinski (2000), overlaps with the definition of stability: the reader may refer to the appendix for more details on the notation. The same constraints can be used to formulate the linear programming that finds an efficient allocation among stable ones.

The previous LP does not consider the total match quality. Consequently, one may end up sacrificing much of the aggregate educational opportunity for the sake of fairness. To limit this trade off, one can impose an additional constraint of the sort

$$\sum_{i \in I} U(i,s) x_{(i,s)} \geq y$$

for some positive number y. Then, a grid search on the values of y can help the evaluator in choosing the right balance between the otherwise incompatible efficiency and fairness requirements.

One of the main drawbacks of using linear programming is that it can be difficult to understand and interpret the solution ex post. Linear programming is often considered a "black box" approach, as it can be challenging to understand how the solution was obtained and what specific constraints or factors led to the final outcome. In some circumstances, the process of how an allocation is obtained is just as important as the outcome itself, and being able to understand and reconstruct all the steps that led to the outcome is crucial. The following section relies on a standard tool in welfare analysis, namely a social welfare function, to identify a deterministic procedure to obtain an opportunity egalitarian allocation.

4 Normative approach

The previous section discussed the trade-off between efficiency and fairness and how the linear programming solution may be seen as unsatisfactory in some cases. In this section, a different approach is proposed that weakens the efficiency and fairness requirements to a point where they can be both considered in a single social welfare function. This approach aims to find a balance between the two requirements and allows for a more comprehensive evaluation of the final outcome.

We begin by characterizing a family of Social Welfare Functions (SWFs), that is assumed to represent the preferences of a central authority, and is used to evaluate stable matchings. As before, we impose stability to be the minimal requirement for a desirable school allocation. Therefore, the central authority will aim at maximizing her SWF over the set of all stable assignments.

Denoting with M the set of stable matchings, $W : M \rightarrow \mathbb{R}$ is the functions that measure the social welfare associated to a stable school assignment. In line with the preferences of an opportunity egalitarian social evaluator, we assume SWFs to satisfy the following axioms.

Monotonicity (MON) - For all $\mu, \mu' \in M$, if $U(i, \mu(i)) \geq U(i, \mu'(i))$ for all $i \in I$, with at least one strict inequality, then $W(\mu) \geq W(\mu')$.

Monotonicity is inspired by Efficiency. This property simply states that improving match quality of someone, without reducing it for anyone else, cannot worsen social welfare.

Additivity (ADD) - For all $\mu \in M$, there exist twice differentiable (almost everywhere) functions $\phi_i : \mathbb{R} \rightarrow \mathbb{R}$, for all $i \in I$, such that $W(\mu) \equiv \sum_{i \in I} \phi_i(U(i, \mu(i)))$.

The previous property imposes our social evaluation to be based on a standard utilitarian aggregation of match qualities. This axioms imposes separability across individuals, so that the comparison between two alternative allocations can be performed by comparing the changes in match quality of the sole individuals with a different matching. This allows for a more straightforward and transparent evaluation of the different allocations and makes it easier to understand how changes in match quality affect the overall matching.

Within type symmetry (SYM) - For all $\mu, \mu' \in M$, if there exists $t \in T$ and $i, i' \in t'$ such that $U(i', \mu'(i')) = U(i, \mu(i)) > U(i', \mu(i')) = U(i, \mu'(i))$, then $W(\mu) = W(\mu')$.

Within type symmetry implements the well-known anonymity principle, adapting it to a context in which circumstances do matter in the social evaluation. In words, this property requires social evaluation not to change if there is a permutation of match qualities involving individuals with the same characteristics. In other words, within a type, it does not matter who has a given match quality level.

Within types inequality neutrality (WTIN) - For all $\mu, \mu' \in M$, if there exist $t \in T$, $i, i' \in t$ and a positive real number δ such that $U(i, \mu'(i)) \geq U(i', \mu'(i'))$, $U(i, \mu(i)) = U(i, \mu'(i)) + \delta$, $U(i', \mu(i')) = U(i', \mu'(i')) - \delta$, and $U(i'', \mu(i'')) = U(i'', \mu'(i''))$ for all other $i'' \in I$, then $W(\mu) = W(\mu')$.

This property strengthen the idea that the match quality distribution within a type is matter of indifference for our social evaluation. Indeed, among individuals with the same circumstances, the final allocation should depend on their preferences, together with school priorities. Although one may argue that this is not the case for the latter, the former are expression of individual freedom which we intend to respect where possible.

Between types inequality aversion (BTIA) - For all $\mu, \mu' \in M$, if there exist two types $t, t' \in T$ such that $t \lhd t'$, $i \in t$, $i' \in t'$ and a positive real number δ such that $U(i', \mu(i')) = U(i', \mu'(i')) - \delta$, $U(i, \mu(i)) = U(i, \mu'(i)) + \delta$ and $U(i'', \mu(i'')) = U(i'', \mu'(i''))$ for all other $i'' \in I$, then $W(\mu) > W(\mu')$.

Between type inequality aversion is an alternative way of expressing the desire for Fairness. In particular, while the Fairness requirement in the previous section abstracts from any type order, BTIA implements the desire of reducing the inequalities between types, allowing for the possibility of using (more favorable) educational outcomes as a compensation for disadvantages in other aspects of life.

As also shown in Peragine (2004), the five axioms above characterize a family of linearly additive social welfare functions. Let ϕ' and ϕ'' denote respectively the first and second derivative of ϕ.

Lemma 2 (Peragine (2004))**.** For all $\mu \in M$, W satisfies MON, ADD, SYM, WTIN and BTIA if and only if

$$W(\mu) = \sum_{i \in I} \phi_i \left(U(i, \mu(i)) \right) \tag{1}$$

where the functions $\phi_i : \mathbb{R} \to \mathbb{R}$ satisfies the following conditions: (i) for all $i \in I$, $\phi_i'' = 0$; (ii) for all $i, i' \in t$, $\phi_i' = \phi_{i'}' > 0$; and (iii) for all $i \in t$, $i' \in t'$ and $t, t' \in T$, if $t \lhd t'$, then $\phi_i' > \phi_{i'}' > 0$.

In the rest of this section, after introducing some necessary definitions and notation, we describe the Stable Opportunity Egalitarian (SOE) algorithm, which identifies the stable allocation that maximizes a given instance of eq. (1).

4.1 The Stable Opportunity Egalitarian algorithm

Irving *et al.* (1987) designed an algorithm to find a stable matching, in one-to-one matching setting, that maximizes preference satisfaction in both sides of the markets. Their algorithm relies on the concepts of "rotations" and "rotation poset" previously introduced by Irving & Leather (1986). Bansal *et al.* (2007) generalizes the concept of rotations to many-to-many matching markets. Cheng *et al.* (2008) show that, under particular restriction on what they call happiness measures, Irving *et al.* (1987) algorithm can be generalized to many-to-one matching setting. In our setting, we draw a parallel between Cheng *et al.* (2008) happiness function and our match quality measure, and use the proposed algorithm to maximize social welfare.

To properly define the algorithm, we need to introduce a series of technical concepts from the matching literature. The following definition formalizes the well known Deferred Acceptance algorithm (Gale & Shapley, 1962) in two different version: one in which students propose to schools and schools choose who to admit, and one in which schools propose to students who then choose where to enroll.

Definition 2 (Deferred Acceptance)**.** (a) We call Student-proposing Deferred Acceptance the following algorithm, and denote μ_I the resulting school allocation. (Step 1)

Each student applies to her first ranked school. Each school collects the applications and temporarily assigns seats to applicants, one at a time, following the priority order up until the quota. Any remaining applicant is rejected. (Step k) Each student who was rejected in the previous step applies to her next most preferred school. Each school considers all the applicants - the new and the temporarily assigned ones from the previous step - and temporarily assigns its seats one at a time following the priority order up until the quota. Any remaining applicant is rejected. (Termination) If in the previous step no student was rejected, then the algorithm terminates and all the temporary assignments become final. (b) The same algorithm is called School-proposing Deferred Acceptance when schools and students are inverted in the process. We denote μ_S the resulting allocation.

Another necessary concept is preferences and priorities reduction, or "pruning." This process helps us, for each student, to keep all possible schools that he can match in a stable matching and remove all the other schools from his preferences. Likewise, for each school, we keep only all possible students that this school can match in a stable matching and remove all the other students from its priority ranking. The first step of pruning is to remove all students who are not matched to school s in the student optimal stable matching from the priority ranking of the school s if s is not filling its capacity. Likewise, we remove all the other schools except the school s from the preferences of the student i who is matched to the missing capacity school s in the student optimal stable matching. This step is due to a well known result in matching theory literature, called *Rural hospitals theorem* (Roth, 1986, 1984). This theorem states that if a school is not filling its capacity in some stable matching, then it will not fill its capacity in any other stable matching. Moreover, this school will get always the same set of students in any stable matching. The second step of pruning is removing all the schools which are ranked above $\mu_I(i)$ in the preferences of student i and the schools which are ranked below $\mu_S(i)$. Since μ_I is the Pareto optimal allocation for students among all stable matchings, there is no stable matching μ where $\mu(i) P_i \mu_I(i)$. Conversely, for any student i, his assignment in any other stable matching $\mu(i)$ is weakly better than $\mu_S(i)$. Hence, for any student i, there is no stable matching μ where $\mu_S(i) P_i \mu(i)$. The third step of pruning is for the priority rankings of schools. The structure of the set of stable matchings holds symmetrically in terms of school's priority ranking, so that μ_S (resp. μ_I) is the best (resp. worst) stable matching from schools' perspective. The final step of pruning is to remove all non mutually acceptable student-school pairs on the priority rankings and the preferences. The formal definition of pruning is the following.

Definition 3 (Pruning)**.** The pruning of preferences and priority rankings in an instance $\mathcal{I}$ is obtained with the following procedure.

1. For any school $s \in S$ such that $|\mu_I(s)| < q_s$: (i) remove from $\succ_s$ all students $i \notin \mu_I(s)$; (ii) for all $i \in \mu_I(s)$ remove from P_i all schools $s' \neq s$.

2. For any student $i \in I$, remove from P_i all schools $s \in S$ such that $s P_i \mu_I(i)$ or $\mu_S(i) P_i s$.

3. For any school $s \in S$, remove from $\succ_s$ all students $i \in I$ such that $i \succ_s i'$ for all $i' \in \mu_S(s)$, or $i'' \succ_s i$, for all $i'' \in \mu_I(s)$.

4. For each $s \in S$ and $i \in I$: (i) if $i \in \succ_s$ but $s \notin P_i$, remove i from $\succ_s$; (ii) if $s \in P_i$ but $i \notin \succ_s$, remove s from P_i.

The pair of pruned preferences and priorities, denoted $(P^\star, \succ^\star)$ define the set of admissible student-school pairs in a stable matching.

We call *graph* of a matching μ, the directed graph $G_\mu = (V_\mu, E_\mu)$ in which: (i) the set of vertices, V_μ, is formed by the set of student-school pairs, $(i, s) \in I \times S$, such that i is the worst student for school s according to $\succ_s^\star$ and $\mu(i) = s$, and (ii) E_μ is a set of oriented edges such that: there is an edge from a vertex (i, s) to a vertex (i', s') whenever s' is the second best school after $\mu(i)$, according to $P_i^\star$. Endowed with the notion of graph of a matching, we can define the following key ingredient of our algorithm.

Definition 4 (Rotations)**.** Let μ denote a school assignment. A *rotation*

$$r = \langle (i_1, s_1), (i_2, s_2), ..., (i_n, s_n) \rangle$$

on $G_\mu = (V_\mu, E_\mu)$, or *exposed rotation* in μ, is a sequence of vertices in V_μ such that for all $j \in \{1, ..., n\}$, there is an oriented edge from (i_j, s_j) to (i_{j+1}, s_{j+1}) where j is taken modulo n. [6] We also call r the exposed rotation in μ starting from i_1.

In words, a rotation on a graph is a cycle that, starting from a given vertex, follows a sequence of edges until it reaches the starting point.

Let $V_\mu \setminus r$ denote the set of student-school pairs which are not in r.

Definition 5 (Eliminating a rotation)**.** A matching μ' *eliminates* an exposed rotation r in μ - denoted $\mu' = \mu \backslash r$ - if, for all student-school pairs in $r = \langle (i_1, s_1), (i_2, s_2), ..., (i_n, s_n) \rangle$, $\mu'(i_j) = s_{j+1}$ where j is taken modulo n, and for all student-school pairs in $V_\mu \setminus r$, $\mu(i) = \mu'(i)$.

In words, a rotation defines a sequence of re-allocations in which student j (who was the least preferred student by the school she is assigned to) goes to his second most preferred school, taking the spot of $j + 1$ (who was the least preferred student by the school he is assigned to) who goes to his second most preferred school and takes the spot of $j + 2$ who then goes to his second most preferred and so on. If we implement this sequence of transfers, we obtain a new assignment which is said to be the matching that eliminates that rotation.

Eliminating an exposed rotation r in μ creates a new assignment $\mu' = \mu \setminus r$. Starting from a student i such that $\mu'(i) \neq \mu_S(i)$, we can expose new rotations in μ'. This process of eliminating an exposed rotation, and exposing new rotations, allows us to define the set of all exposed rotations, starting from a given matching μ. Since some rotations will be exposed only after eliminating others, it is useful to talk about successors and

[6] That is, if $j = n$, then $j + 1 = 1$.

predecessors of a rotation. We say that a rotation ρ is a successor of r - denoted $r < \rho$ - if ρ is an exposed rotation in $\mu \setminus r$ and it is not possible to expose ρ without eliminating r. Indeed, since there may be multiple exposed rotations in an assignment, a rotation ρ may be only exposed after eliminating multiple rotations.

For each rotation $\rho = \langle (i_1, s_1), (i_2, s_2), ..., (i_n, s_n) \rangle$ exposed on μ, we define a *weight* $\omega(\rho) = W(\mu \setminus \rho) - W(\mu)$. Observe that, because match quality satisfies IND and W is linearly additive, $\omega(\rho) = \sum_{i_j \in \rho} \phi_{i_j}(U(i_j, s_{j+1})) - \phi_{i_j}(U(i_j, s_j))$ for j taken modulo n.

Given an instance $\mathcal{I} = \langle I, S, P, \succ, q, T, \lhd, U \rangle$, denote $\Pi(\mathcal{I}, \mu)$ the set of all exposed rotations starting from μ. Notice that the relation $<$ defines a partial order on $\Pi(\mathcal{I}, \mu)$, so that $(\Pi(\mathcal{I}, \mu), <)$ forms a partially ordered set. A closed set in the poset $\Pi(\mathcal{I}, \mu)$ is a subset $C(\mathcal{I}, \mu)$ of $\Pi(\mathcal{I}, \mu)$ such that

$$\rho \in C(\mathcal{I}, \mu), r < \rho \Rightarrow r \in C(\mathcal{I}, \mu)$$

In words, a subset of rotations $C(\mathcal{I}, \mu) \subset \Pi(\mathcal{I}, \mu)$ is closed if it contains all the predecessors of its elements. Following Bansal *et al.* (2007), we can formalize the following result.

Lemma 3 (Bansal *et al.* (2007))**.** Let $\mathcal{I}$ be an instance of school choice problem with EOp components and μ_I its student-proposing DA assignment. There is a one-to-one correspondence between the closed subsets of $\Pi(\mathcal{I}, \mu_I)$ and the set of all stable matchings of $\mathcal{I}$: each closed subset $C(\mathcal{I}, \mu_I)$ of $\Pi(\mathcal{I}, \mu_I)$ corresponds to a unique stable matching generated by eliminating all the rotations in $C(\mathcal{I}, \mu_I)$.

The previous lemma provides a powerful result which tells us that we can somehow explore the set of all stable matchings via looking at all exposed rotations, starting from the student-proposing DA assignment. While listing all the stable matchings of $\mathcal{I}$ is NP-hard, $\Pi(\mathcal{I}, \mu_I)$ can be constructed with an efficient algorithm[7] (see Cheng *et al.*, 2008) in the following way.

Step 0: Run student proposing and school proposing DAs, and find the respective matchings, μ_I and μ_S.

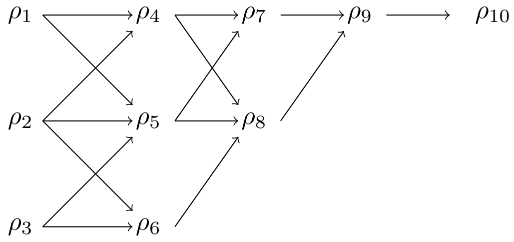
Step 1: (1.1.) Prune preferences and priority rankings as in Definition 3. (1.2.) Form the graph of matching μ_I - G_{μ_I}. Starting from a student i who is not matched to $\mu_S(i)$, find an exposed rotation ρ_1. (1.2.) Add this rotation to the rotation poset $(\Pi(\mathcal{I}, \mu_I))$ and compute $\omega(\rho_1) = W(\mu_I \setminus \rho_1) - W(\mu_I)$.

Step 2: (2.1.) Substitute $\mu_I \setminus \rho_1$ to μ_I in Definition 3 and prune preferences and priority rankings. (2.2.) Form the graph of $\mu_I \setminus \rho_1$ and, starting from a student i who is not matched to $\mu_S(i)$ find an exposed rotation ρ_2. (2.2.) Add this rotation to the rotation poset and compute its weight $\omega(\rho_2) = W(\mu_I \setminus \{\rho_1, \rho_2\}) - W(\mu_I \setminus \rho_1)$.

Step k: (k.1.) Prune based on $\mu_I \setminus \{\rho_1, ...\rho_{k-1}\}$ and μ_S. (k.2.) Form the graph of $\mu_I \setminus \{\rho_1, ...\rho_{k-1}\}$ and expose a rotation ρ_k. (k.3.) Add this rotation in the rotation

[7] An example of this algorithm is reported in the Appendix.

Figure 1: A graphical example of a rotation poset.



poset and compute its weight.

Termination: Terminate this procedure until, after eliminating a rotation, you obtain μ_S.

Notice that, at a given assignment μ, a student can belong to at most one of its exposed rotations. Moreover, the pruning process shrinks the number of possible rotation that can be exposed. In addition, no student-school pair can belong to more than one rotation (Cheng *et al.*, 2008). Consequently, there are at most $|I| \cdot |S|$ rotations. After finding all the rotations, we need to choose a closed subset of $\Pi(\mathcal{I}, \mu_I)$ such that eliminating all the rotations in this subset will maximize our social evaluation function. A graphical example of a rotation poset is provided in Figure 1. The graph in is composed of vertices, which correspond to the rotations, and edges that connect each rotation to its successors. For a more detailed explanation on how to construct the such a graph, see the example in Appendix.

Finding the maximum closed subset of $\Pi(\mathcal{I}, \mu_I)$ is a selection problem like the one introduced by Balinski (1970) and Rhys (1970). Picard (1976) shows that this problem can be easily solved using Linear Programming. In particular, let n be the number of rotations in $\Pi(\mathcal{I}, \mu_I)$, then finding the maximum closed subset is equivalent to finding the vector $\mathbf{x} \in \{0, 1\}^n$ that solves

$$\max_{\mathbf{x}} z = \sum_{j=1}^{n} \omega\left(\rho_j\right) x_j + \lambda \sum_{j=1}^{n} \sum_{h=1}^{n} a_{jh} x_j \left(-1 + x_h\right) \tag{2}$$

where $a_{jh} = 1$ if, on the rotation graph, there is a directed edge from ρ_h to ρ_j (i. e. ρ_j is a successor of ρ_h) and $a_{jh} = 0$ otherwise, and λ must be an arbitrarily large real number.[8] In words, $x_j = 1$ means that the rotation j is in the subset. For each possible subset of $\Pi(\mathcal{I}, \mu_I)$, the first element of Eq. 2 sums the weights of the rotations in the subset and the second element checks that, for each of these rotations, the respective predecessors are included in the subset. Suppose that ρ_h is a predecessor of ρ_j, ρ_j is in the subset but ρ_h is not, then $a_{jh} x_j \left(-1 + x_h\right) = 1 \cdot 1(-1 + 0) = -1$.

Picard (1976) shows that there is a more efficient way of solving this problem by reducing the graph of the rotation poset to a network flow graph and applying a min-

[8] It must be $\lambda > \max\{\omega(\rho_j)\}_{j=1}^{n}$ to ensure that the optimal solution satisfies $\sum_{j=1}^{n} \sum_{h=1}^{n} a_{jh} x_j \left(-1 + x_h\right) = 0$, which means that we have found a closed subset.

cut algorithm. This is the procedure we suggest in our algorithm. However, to maintain a simple exposition, we do not include the details of this methodology in the main text, the interested reader may refer to the Appendix.

We can now introduce the Stable Opportunity Egalitarian (SOE) algorithm, which is structured as follows. First, set a functional form for W. Second, construct the rotation poset. Third, construct network flow of rotations. Fourth, apply an efficient minimum cut algorithm to identify the closed subset of rotation we need to eliminate to maximise W. Fifth, return the optimal allocation. The pseudo-code of the SOE algorithm is as follows.

Algorithm 1 The SOE algorithm (Cheng *et al.*, 2008)

Input: W and $\mathcal{I} = \langle I, S, P, \succ, q, T, \trianglelefteq, U \rangle$
Compute μ_I and μ_S.
Prune preferences of students and priority rankings of schools;
$\mu \leftarrow \mu_I$, $\Pi(\mathcal{I}, \mu) = \emptyset$
while $\mu \neq \mu_S$ **do**
    Find $i^* \in I$ such that $\mu(i^*) \neq \mu_S(i^*)$
    Find a rotation ρ exposed in μ starting from i^*
    Add ρ to $\Pi(\mathcal{I}, \mu)$
    $\omega(\rho) \leftarrow W(\mu \setminus \rho) - W(\mu)$
    $\mu \leftarrow \mu \setminus \rho$
end while
Construct the network flow of rotations and use min-cut to find $C^* = \arg\max_{C \subset \Pi(\mathcal{I},\mu)} \sum_{r \in C} \omega(r)$;
$\mu \leftarrow \mu_I \setminus C^*$
Return μ

Let us recall that there are at most $|I| \cdot |S|$ rotations. Let t represent the longest running time that takes to calculate $\omega(\rho)$ for some rotation ρ. It takes at most $|I|^4$ running time to find the optimal closed subset of $\Pi(\mathcal{I}, \mu_I)$ (see Gusfield & Irving (1989), page 129-133). Putting all these calculations together, we can conclude that the SOE algorithm takes at most $|I||S|t + |I|^4$ running time; this is an efficient running time performance.

The following two statements prove the correctness of the SOE algorithm.

Lemma 4. Let $\mathcal{I}$ be a school choice instance. Let C be a closed subset of $\Pi(\mathcal{I}, \mu_I)$, and μ be the stable matching obtained by eliminating all the rotations in C. If $U : I \times S \to \mathbb{R}$ satisfies IND, and W satisfies MON, ADD, WTIN, SYM and BTIA, then $W(\mu) = W(\mu_I) + \sum_{r \in C} \omega(r)$.

The previous result is a direct implication of the additive decomposability of Eq. 1, combined with the independence property satisfied by the match quality measures in our setting. We now provide the main result of the paper.

Theorem 1. Let $\mathcal{I} = \langle I, S, P, \succ, q, T, \triangleleft, U \rangle$ be an instance of centralized school choice setting with EOp components. Let $W : M \to \mathbb{R}$ be a social welfare function that satisfies MON, ADD, WTIN, SYM and BTIA. Then, the Stable Opportunity Egalitarian algorithm maximizes W over the set of stable assignments for $\mathcal{I}$.

Proof. By Lemma 4, finding a closed subset C that maximizes W is equivalent to finding a matching that maximizes our social evaluation function. By Lemma 3, this matching will be stable. □

5 Discussion

In this section we comment on relevant features of the SOE algorithm, and its link with the normative principles behind the EOp paradigm.

Let us begin by recalling that Student-proposing DA delivers the best allocation, among stable ones, in terms of preference satisfaction. Consequently, if match quality corresponds to student's preference satisfaction - U_P from Definition 1 - then μ_I is optimal, for any SWF as in eq. (1). The reader may notice that with an U_P type of match quality, it is as if the central planner aligns his preferences to those of the students. In the EOp paradigm, this situation corresponds to a strong reward principle which calls for preservation of all those inequalities due to factors within individual control: preferences in our context. To put it differently, the use of Student-proposing DA to allocate students can be justified by the normative principle of respecting students' preferences and holding individuals responsible for them (Dworkin, 1981b). Notice that, although we have been focusing on EOp between students, one can formulate the symmetric problem of equalizing school's opportunities for 'good' students.

The SOE algorithm deviates from either μ_I or μ_S whenever the central authority's preferences do not coincide with those of the students or schools. This is the case, for example, if the evaluator does not want to hold individuals responsible for their preferences, which can be influenced by factors out of individual control. In the matching literature, affirmative action policies are directly or indirectly aimed at modifying priority rankings in favour of particular groups. In our context, we achieve a similar result by shifting the attention of the social evaluator toward other preferences of which the match quality measure, in combination with the social welfare function, offer a representation.

Equalizing opportunity is tightly linked with the rationale behind affirmative action policies. Intuitively, if we want to improve chances for disadvantaged students to get into preferred schools, it may be sufficient to modify the schools' priority rankings by upward moving these students. Observe that we can still focus on multiple groups at the same time by calibrating the *bonuses* given to students from different types. In line with the principle of respect for students' preferences, we can run Student-proposing DA and implement the resulting allocation, call it $\mu*$. Despite the adverse effect this can have in theory (see Kojima, 2012), we should expect this to improve the situation for the disadvantaged groups.

The matching $\mu*$ may fail to satisfy the Stability requirement which is defined in terms of the original preferences and priority rankings. However, if $\mu*$ satisfies Stability, then it is possible to define a match quality measure U and a social welfare function W such that the SOE algorithm identifies $\mu*$ as the optimal allocation. Under standard affirmative action policies, it is often necessary to weaken the Stability condition for it to be satisfied. Our exercise points out a way of implementing affirmative action policies under the standard notion of Stability, and the EOp paradigm identifies normative reasons for those policy not to be desirable: the (dis)alignment between students' and social evaluator's preferences.

The SOE algorithm has the advantage of being a deterministic procedure in which we can reconstruct all the passages that lead to a given allocation. Intuitively, the starting point of the algorithm is always the best possible allocation for students. Then, it proceeds with sacrificing student's preference satisfaction in order to maximize the evaluator's preferences, represented by the social welfare function. In this sense, when moving away from Student-proposing DA, the algorithm trades off Pareto efficiency and maximization of the evaluator's preferences. It is worth underlining here that, in line with our discussion in Section 3, we can use linear programming to maximize W under the stability constraints. Abstracting here form all issues related to the computational complexity of the two procedures, the SOE algorithm is clearly more transparent and easier to back-track. For sensitive matters like school allocation, we believe this constitutes a strong motivation for preferring our deterministic procedure to a linear programming solution.

We conclude this section by discussing other applications of the SOE algorithm. Problems like doctor-hospital or refugees allocation are other instances of many-to-one matching problems in which opportunity egalitarian principles find application. In our setting we focus on the problem of equalizing student's opportunity for good education. Social planners may be concerned with the symmetric problem of equalizing hospital's opportunity for good doctors, with the aim of reducing regional disparities in the health care system. The SOE algorithm offers a solution to this problem as well. Clearly, the same holds for any many-to-one matching setting in which can be expressed, as we do in this paper, as a problem of maximization of a linear social welfare function over the set of stable matchings.

6 Conclusion

We have investigated the possibility of respecting equality of opportunity principles in centralized school choice. We have shown the incompatibility between out new Fairness principle and two other desirable properties from the literature. We have linear programming solutions to identify fair allocations that satisfy standard stability conditions but argued against the use of linear programming for this allocation problem. This motivated the implementation of a normative approach, typical of the social choice literature, to evaluate the desirability of a stable allocation. We defined a family of opportunity egalitarian social welfare functions and described a deterministic procedure - the SOE algorithm - to maximize social welfare over the set of stable matchings.

To the best of our knowledge, this is the first paper to draw a clear connection between the school choice and the equality of opportunity literature, and we believe this to be only a first step toward a more consistent dialogue between those two literatures. There is great scope for further exploring how existing algorithm in the matching literature can be used to solve complex fairness issues, like equality of opportunity, that go beyond the standard aversion to inequality. The literature has proposed different families of opportunity egalitarian social welfare functions, many of which fail to satisfy the independence property. Identifying algorithmic solutions to the problem of maximizing non-linear opportunity egalitarian social welfare functions over the set of stable matchings is highly ranked in our research agenda.

Finally, we underline that our algorithmic solution is applicable to other many-to-one matching problems like the allocation of new doctors to hospitals or refugees to hosting centers. In all these settings have been extensively explored by the matching literature, and the opportunity egalitarian principle offers new ways of addressing and assessing this important allocation problems.

References

Abdulkadiroğlu, Atila, & Sönmez, Tayfun. 2003. School choice: A mechanism design approach. *American economic review*, **93**(3), 729–747.

Abdulkadiroğlu, Atila, Pathak, Parag A, Schellenberg, Jonathan, & Walters, Christopher R. 2020. Do parents value school effectiveness? *American Economic Review*, **110**(5), 1502–39.

Abdulkadiroglu, Atila, Dur, Umut M, & Grigoryan, Aram. 2021. *School Assignment by Match Quality*. Tech. rept. National Bureau of Economic Research.

Afacan, Mustafa Oğuz, & Salman, Umutcan. 2016. Affirmative actions: The Boston mechanism case. *Economics Letters*, **141**, 95–97.

Alesina, Alberto, Stantcheva, Stefanie, & Teso, Edoardo. 2017. *Intergenerational mobility and support for redistribution*. Tech. rept. National Bureau of Economic Research.

Atkinson, Anthony B, & Bourguignon, Francois. 1982. The comparison of multi-dimensioned distributions of economic status. *The Review of Economic Studies*, **49**(2), 183–201.

Baïou, Mourad, & Balinski, Michel. 2000. The stable admissions polytope. *Mathematical programming*, **87**(3), 427–439.

Balinski, Michel L. 1970. On a selection problem. *Management Science*, **17**(3), 230–231.

Bansal, Vipul, Agrawal, Aseem, & Malhotra, Varun S. 2007. Polynomial time algorithm for an optimal stable assignment with multiple partners. *Theoretical Computer Science*, **379**(3), 317–328.

Baïou, Mourad, & Balinski, Michel. 2000. The stable admissions polytope. *Mathematical Programming*, **87**, 427–439.

Bossert, Walter. 1995. Redistribution mechanisms based on individual characteristics. *Mathematical Social Sciences*, **29**(1), 1–17.

Cappelen, Alexander W, Hole, Astri Drange, Sørensen, Erik Ø, & Tungodden, Bertil. 2007. The pluralism of fairness ideals: An experimental approach. *American Economic Review*, **97**(3), 818–827.

Cappelen, Alexander W, Konow, James, Sørensen, Erik Ø, & Tungodden, Bertil. 2013. Just luck: An experimental study of risk-taking and fairness. *American Economic Review*, **103**(4), 1398–1413.

Cheng, Christine, McDermid, Eric, & Suzuki, Ichiro. 2008. A unified approach to finding good stable matchings in the hospitals/residents setting. *Theoretical Computer Science*, **400**(1-3), 84–99.

Corak, Miles. 2013. Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, **27**(3), 79–102.

Doğan, Battal. 2016. Responsive affirmative action in school choice. *Journal of Economic Theory*, **165**, 69–105.

Dworkin, Ronald. 1981a. What is equality? Part 1: Equality of welfare. *Philosophy & public affairs*, 185–246.

Dworkin, Ronald. 1981b. What is equality? Part 2: Equality of resources. *Philosophy & public affairs*, 283–345.

Fleurbaey, Marc. 1994. On fair compensation. *Theory and decision*, **36**(3), 277–307.

Fleurbaey, Marc, & Peragine, Vito. 2013. Ex ante versus ex post equality of opportunity. *Economica*, **80**(317), 118–130.

Ford, Lester Randolph, & Fulkerson, Delbert R. 1956. Maximal flow through a network. *Canadian journal of Mathematics*, **8**, 399–404.

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

Gusfield, Dan, & Irving, Robert W. 1989. *The stable marriage problem: structure and algorithms*. MIT press.

Hafalir, Isa E, Yenmez, M Bumin, & Yildirim, Muhammed A. 2013. Effective affirmative action in school choice. *Theoretical Economics*, **8**(2), 325–363.

Irving, Robert W, & Leather, Paul. 1986. The complexity of counting stable marriages. *SIAM Journal on Computing*, **15**(3), 655–667.

Irving, Robert W, Leather, Paul, & Gusfield, Dan. 1987. An efficient algorithm for the "optimal" stable marriage. *Journal of the ACM (JACM)*, **34**(3), 532–543.

Kojima, Fuhito. 2012. School choice: Impossibilities for affirmative action. *Games and Economic Behavior*, **75**(2), 685–693.

Marrero, Gustavo A, & Rodríguez, Juan G. 2013. Inequality of opportunity and growth. *Journal of development Economics*, **104**, 107–122.

Peragine, Vito. 2004. Ranking income distributions according to equality of opportunity. *The Journal of Economic Inequality*, **2**(1), 11–30.

Peragine, Vitorocco. 2002. Opportunity egalitarianism and income inequality. *Mathematical Social Sciences*, **44**(1), 45–64.

Picard, Jean-Claude. 1976. Maximal closure of a graph and applications to combinatorial problems. *Management science*, **22**(11), 1268–1272.

Ramos, Xavier, & Van de Gaer, Dirk. 2016. Approaches to inequality of opportunity: Principles, measures and evidence. *Journal of Economic Surveys*, **30**(5), 855–883.

Rawls, John. 1971. *A theory of justice.* Harvard Press, Cambridge.

Rhys, John MW. 1970. A selection problem of shared fixed costs and network flows. *Management Science*, **17**(3), 200–207.

Roemer, John E. 1993. A pragmatic theory of responsibility for the egalitarian planner. *Philosophy & Public Affairs*, 146–166.

Roemer, John E. 1998. *Theories of distributive justice.* Harvard University Press.

Roemer, John E, & Trannoy, Alain. 2015. Equality of opportunity. *Pages 217–300 of: Handbook of income distribution*, vol. 2. Elsevier.

Roemer, John E, & Ünveren, Burak. 2017. Dynamic equality of opportunity. *Economica*, **84**(334), 322–343.

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**(6), 991–1016.

Roth, Alvin E. 1986. On the allocation of residents to rural hospitals: a general property of two-sided matching markets. *Econometrica: Journal of the Econometric Society*, 425–427.

Sen, Amartya, *et al.* 1980. Equality of what?

Appendix A: Linear programming for a fair stable allocation

In this appendix, we provide additional details on the linear programming formulation to find the Fair stable allocation. This formulation is provided by Baïou & Balinski (2000).

A school choice graph $G = (V, E)$ is a directed graph. The set of vertices is composed of mutually acceptable school-student pairs; in our case $V = I \times S$. There is an edge from $v = (i, s)$ to $v' = (i', s')$, $e = (v, v') \in E$, if either $i = i'$ and $s' P_i s$ or if $s = s'$ and $i' \succ_s i$. The following example represents a school choice graph of an instance: $I = \{i_1, i_2, i_3\}$, $S = \{s_1, s_2, s_3\}$, with preferences and priority rankings

P_{i_1}	P_{i_2}	P_{i_3}	s_1	s_2	s_3
s_3	s_1	s_1	i_3	i_1	i_1
s_2	s_3	s_2	i_2	i_3	i_2
s_1	s_2	s_3	i_1	i_2	i_3

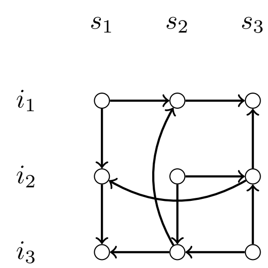


Figure 2: The graph $G = (V, E)$ of a school choice problem

On the graph above, the arcs implied by transitivity are omitted. For example, on the graph, it is obvious that $s_3 P_{i_1} s_1$.

A vertex v' is a successor of another vertex v if there is an edge from v to v', i.e. $e = (v, v') \in E$. For example on the graph above, (s_2, i_1) and (s_1, i_2) are the successors of (s_1, i_1). Let V^+ be the set of all vertices $v = (s, i) \in V$ that have at least $q_s - 1$ successors such that for each successor $v' = (s', i')$ of $v = (s, i)$, $s = s'$. V^- is the set of all other vertices, i.e. $V^- = V \setminus V^+$.

For each school $s \in S$, a *shaft* of $s \in S$ consists of a vertex $(i, s) \in V^+$ which is the base, and all of its successors in V^+. For each school, there may exist multiple student-school pairs in V^+ but only one shaft, which we denote $\mathcal{S}(i, s)$ to highlight that $(i, s) \in V^+$ is the base.

A *tooth* $\mathcal{T}$ of $(i, s) \in V$ - written $\mathcal{T}(i, s)$ - consists of the vertex (i, s) which is the source, and all of its successors (i', s') such that $i = i'$.

A *comb* $\mathcal{C}$ of $s \in S$ is the union of its shaft $\mathcal{S}(i, s)$ and exactly q_s teeth of $(i', s) \in \mathcal{S}(i, s)$, including $\mathcal{T}(i, s)$. It is written as $\mathcal{C}(i, s)$. The teeth are decided by the priority ranking of s. Hence if $\mathcal{T}(i'', s) \in \mathcal{C}(i, s)$, then any teeth $\mathcal{T}(i', s)$ such that $i' \succ_s i''$ are also in $\mathcal{C}(i, s)$. The set of all combs of s is $\mathbb{C}_s$.

The fourth condition of LP:

$$\sum_{(i',s)\in\mathcal{C}(i,s)} x_{(i',s)} \geq q_s \qquad \forall s \in S, \forall \mathcal{C}(i,s) \in \mathbb{C}_s$$

checks that each comb of s contains at least q_s (s, i) vertices such that $\mu(i) = s$ or $x_{(i,s)} = 1$. As shown in Theorem 3 of Baïou & Balinski (2000), this condition is necessary and sufficient for the assignment to be stable.

Appendix B: Example of rotation poset

In this appendix we adapt an example from Irving *et al.* (1987) to our school choice setting.

Let there be 8 students and 8 schools, each with capacity of one. Let the preferences and priority rankings be as follows.

$$\begin{array}{llll}
i_1: & s_3, s_1, s_5, s_7, s_4, s_2, s_8, s_6 & s_1: & i_4, i_3, i_8, i_1, i_2, i_5, i_7, i_6 \\
i_2: & s_6, s_1, s_3, s_4, s_8, s_7, s_5, s_2 & s_2: & i_3, i_7, i_5, i_8, i_6, i_4, i_1, i_2 \\
i_3: & s_7, s_4, s_3, s_6, s_5, s_1, s_2, s_8 & s_3: & i_7, i_5, i_8, i_3, i_6, i_2, i_1, i_4 \\
i_4: & s_5, s_3, s_8, s_2, s_6, s_1, s_4, s_7 & s_4: & i_6, i_4, i_2, i_7, i_3, i_1, i_5, i_8 \\
i_5: & s_4, s_1, s_2, s_8, s_7, s_3, s_6, s_5 & s_5: & i_8, i_7, i_1, i_5, i_6, i_4, i_3, i_2 \\
i_6: & s_6, s_2, s_5, s_7, s_8, s_4, s_3, s_1 & s_6: & i_5, i_4, i_7, i_6, i_2, i_8, i_3, i_1 \\
i_7: & s_7, s_8, s_1, s_6, s_2, s_3, s_4, s_5 & s_7: & i_1, i_4, i_5, i_6, i_2, i_8, i_3, i_7 \\
i_8: & s_2, s_6, s_7, s_1, s_8, s_3, s_4, s_5 & s_8: & i_2, i_5, i_4, i_3, i_7, i_8, i_1, i_6
\end{array}$$

Student proposing and school proposing DA give:

$$\mu_I: \begin{array}{cccccccc} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ i_2 & i_8 & i_1 & i_5 & i_4 & i_6 & i_3 & i_7 \end{array}$$

$$\mu_S: \begin{array}{cccccccc} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ i_4 & i_3 & i_7 & i_6 & i_8 & i_5 & i_1 & i_2 \end{array}$$

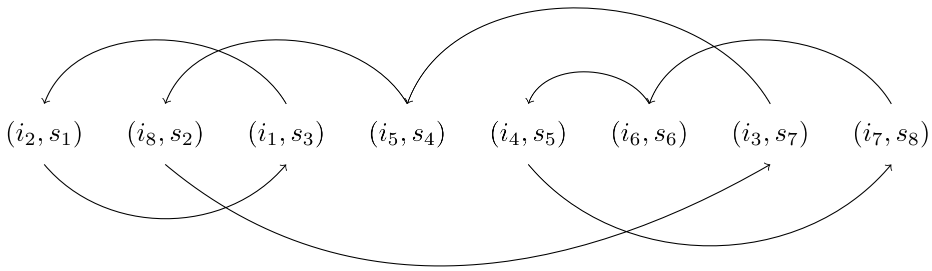
Let us start by pruning preferences and priority rankings. For example, take student 2. He is matched with s_1 in μ_I so there is no stable matching in which he will be assigned to a school he prefers to s_1. Therefore, we can remove s_6 from his preferences. At the same time, $\mu_S(i_2) = s_8$ so there is no stable assignment in which student 2 attends schools 7, 5 or 2, which are all below s_8 in i_2's preference ranking. Hence, the pruned preferences for student 2 do not include schools 6,7,5 and 2. Vice versa, we should also remove student 2 from the priority ranking of these schools.

Consider now school 1. As all the other schools in this instance, it obtains its favourite student under μ_S, so no student should be removed from its ranking according to this criterion. However, under student proposing DA, $\mu_I(s_1) = i_2$ so students 5, 7 and 6, which are all below student 2 in s_1's priority should be removed. Vice versa, we should remove school 1 from the preference list of these students.

By repeating the reasoning above for all students and schools, we get to the following pruned preferences:

$i_1:$	$s_3, s_1, s_5, s_7, \bullet, \bullet, \bullet, \bullet$	$s_1:$	$i_4, i_3, i_8, i_1, i_2, \bullet, \bullet, \bullet$
$i_2:$	$\bullet, s_1, s_3, s_4, s_8, \bullet, \bullet, \bullet$	$s_2:$	$i_3, i_7, i_5, i_8, \bullet, \bullet, \bullet, \bullet$
$i_3:$	$s_7, s_4, s_3, \bullet, \bullet, s_1, s_2, \bullet$	$s_3:$	$i_7, i_5, i_8, i_3, \bullet, i_2, i_1, \bullet$
$i_4:$	$s_5, \bullet, s_8, \bullet, s_6, s_1, \bullet, \bullet$	$s_4:$	$i_6, \bullet, i_2, \bullet, i_3, \bullet, i_5, \bullet$
$i_5:$	$s_4, \bullet, s_2, s_8, s_7, s_3, s_6, \bullet$	$s_5:$	$i_8, \bullet, i_1, \bullet, i_6, i_4, \bullet, \bullet$
$i_6:$	$s_6, \bullet, s_5, s_7, \bullet, s_4, \bullet, \bullet$	$s_6:$	$i_5, i_4, i_7, i_6, \bullet, \bullet, \bullet, \bullet$
$i_7:$	$\bullet, s_8, \bullet, s_6, s_2, s_3, \bullet, \bullet$	$s_7:$	$i_1, \bullet, i_5, i_6, \bullet, i_8, i_3, \bullet$
$i_8:$	$s_2, \bullet, s_7, s_1, \bullet, s_3, \bullet, s_5$	$s_8:$	$i_2, i_5, i_4, \bullet, i_7, \bullet, \bullet, \bullet$

To construct the rotation poset, we start with the graph G_{μ_I} whose vertices are pairs (i, s) such that $\mu_I(i) = s$ and i is the worst student in s's priority ranking. In this example, the vertices coincides with all student-school pairs defined by μ_I. The directed edges are drawn from a vertex (i, s) to a vertex (i', s') where s' follows s in student i's pruned preferences. The graph of student proposing DA is the following.



We can see that three rotations are exposed:
$\rho_1 = \langle (i_2, s_1), (i_1, s_3) \rangle$
$\rho_2 = \langle (i_8, s_2), (i_3, s_7), (i_5, s_4) \rangle$
$\rho_3 = \langle (i_4, s_5), (i_7, s_8), (i_6, s_6) \rangle$
After eliminating ρ_1, we obtain

$$\mu_I \setminus \rho_1 : \begin{array}{cccccccc} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ i_1 & i_8 & i_2 & i_5 & i_4 & i_6 & i_3 & i_7 \end{array}$$

We should now prune again preferences and priority rankings. Remove s_1 (resp. s_3), and any other school above it, in the preferences of i_2 (resp. i_1). Vice versa, remove i_2 (resp. i_1), and any other student below her, in the priority ranking of s_1 (resp. s_3). Finally, check that all remaining pairs are mutually acceptable.

On the graph of $\mu_I \setminus \rho_1$, we still expose ρ_2, ρ_3 but there is no new exposed rotation. Hence, there is no direct successor of ρ_1. If we eliminate ρ_2 we get

$$\mu_I \setminus \{\rho_1, \rho_2\} : \begin{array}{cccccccc} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ i_1 & i_5 & i_2 & i_3 & i_4 & i_6 & i_8 & i_7 \end{array}$$

and on $G_{\mu_I \setminus \{\rho_1, \rho_2\}}$ we expose a new rotation: $\rho_4 = \langle (i_2, s_3), (i_3, s_4) \rangle$. This rotation is a successor of ρ_1 and ρ_2 together.

Let us continue with eliminating ρ_4.

$$\mu_I \setminus \{\rho_1, \rho_2, \rho_4\} : \begin{array}{cccccccc} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ i_1 & i_5 & i_3 & i_2 & i_4 & i_6 & i_8 & i_7 \end{array}$$

After pruning again, observe that on the graph of $\mu_I \setminus \{\rho_1, \rho_2, \rho_4\}$, the only exposed rotation is ρ_3. We should then eliminate it to get:

$$\mu_I \setminus \{\rho_1, \rho_2, \rho_3 \rho_4\} : \begin{array}{cccccccc} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ i_1 & i_5 & i_3 & i_2 & i_6 & i_7 & i_8 & i_4 \end{array}$$

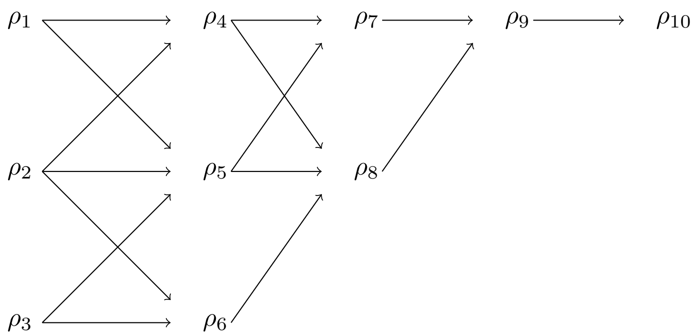
on which graph we expose two rotations:

$\rho_5 = \langle (i_1, s_1), (i_6, s_5), (i_8, s_7) \rangle$

$\rho_6 = \langle (i_5, s_2), (i_4, s_8), (i_7, s_6) \rangle$

Observe that to expose ρ_6, which involves schools 2, 8 and 6 it is not necessary to eliminate ρ_1 which concerns only schools 1 and 3. Therefore, ρ_6 is a successor of ρ_2 and ρ_3 alone. Differently, ρ_5 is a successor of ρ_1, ρ_2, ρ_3. The other rotations that can be exposed after eliminating only ρ_4, ρ_5 or ρ_4, ρ_5, ρ_6 are, respectively, $\rho_7 = \langle (i_8, s_1), (i_3, s_3) \rangle$ and $\rho_8 = \langle (i_2, s_4), (i_5, s_8), (i_6, s_7) \rangle$. The last two rotations are $\rho_9 = \langle (i_8, s_3), (i_1, s_5), (i_5, s_7) \rangle$ and $\rho_{10} = \langle (i_3, s_1), (i_7, s_2), (i_5, s_3), (i_4, s_6) \rangle$. Where ρ_9 is exposed after eliminating ρ_7 and ρ_8, and ρ_{10} is a the successor of ρ_9. After eliminating ρ_{10} we obtain school proposing DA - i. e. $\mu_I \setminus \{\rho_j\}_{j=1}^{10} = \mu_S$.

The consequent rotation poset is



Appendix C: Network flow of rotations

To find the optimal closed subset of rotations based on the method of Picard (1976), define the network flow of rotations. This is a directed weighted graph $\mathcal{G}_\Pi = (\mathcal{V}_\Pi, \mathcal{E}_\Pi)$ where the set of vertices is composed of a source vertex, $\mathcal{B}$, all the rotations and a terminal vertex, $\mathcal{T}$, i.e. $\mathcal{V}_\Pi \equiv \mathcal{B} \cup \Pi(\mathcal{I}, \mu) \cup \mathcal{T}$. There is a directed edge from $\mathcal{B}$ to each $r \in \Pi(\mathcal{I}, \mu)$ if $\omega(r) < 0$ and the weight of the edge is $|\omega(r)|$, i.e. $e = (\mathcal{B}, r) \in \mathcal{E}_\Pi$ if $\omega(r) < 0$ and $w(e) = |\omega(r)|$. There is a directed edge from $\rho \in \Pi(\mathcal{I}, \mu)$ to $r \in \Pi(\mathcal{I}, \mu)$ if r is the successor of ρ and the weight of the edge is infinity, i.e. for $\rho, r \in \Pi(\mathcal{I}, \mu)$, $e = (\rho, r) \in \mathcal{E}_\Pi$ if $\rho < r$ and $w(e) = \infty$. There is directed edge from each rotation $r \in \Pi(\mathcal{I}, \mu)$ to the terminal vertex $\mathcal{T}$ if $\omega(r) > 0$ and the weight of the edge is the weight of the rotation, i.e. $e = (r, \mathcal{T}) \in \mathcal{E}_\Pi$ if $\omega(r) > 0$ and $w(e) = \omega(r)$.

A $(\mathcal{B} - \mathcal{T})$ cut of $\mathcal{G}_\Pi$ is a partitioning of the set of vertices on the flow network into two sets, B and $\overline{B}$, such that $\mathcal{B} \in B$, $\mathcal{T} \in \overline{B}$, $B \cup \overline{B} = \mathcal{V}_\Pi$ and $B \cap \overline{B} = \emptyset$. The capacity of a $(\mathcal{B} - \mathcal{T})$ cut, denoted by $c(B, \overline{B})$, is the sum of the weights of the edges which go from the vertices in B to the vertices in $\overline{B}$. Pictorially, if the partition is a line that divides the graph of the flow network in two, the capacity is given by the sum of the weight of the edges that intersect this line in the direction from the source to the sink. We call minimum $(\mathcal{B} - \mathcal{T})$ the cut with minimal capacity among all possible cuts.

In our flow network, we can think of the rotations (the vertices) as stations that have the capacity of letting a certain flow passing through. This capacity is measured by the absolute value of the weight and the sign of the weight determines the direction the flow is sent toward: either toward the beginning $\mathcal{B}$ or toward the terminal $\mathcal{T}$. A maximum flow algorithm, for example Ford-Fulkerson algorithm (Ford & Fulkerson, 1956), would find the maximum flow this network can carry from the source vertex to the terminal vertex. Because of the duality between maximum flow and minimum cut, the maximum flow can be found with a minimum cut algorithm and this algorithm would give us the set of vertices which will carry the maximum flow. Since the capacity of each vertex in our flow network of rotations is the added value of this rotation to the social welfare function if it is eliminated, a minimum $(\mathcal{B} - \mathcal{T})$ cut in our flow network will give us the set of rotations we need to eliminate, consecutively, to find the stable matching that maximizes social welfare.