The Design of Teacher Assignment: Theory and Evidence*

Julien Combe† Olivier Tercieux‡ Camille Terrier§

April 20, 2015

Preliminary and incomplete. Please do not distribute.

Abstract

In several countries, teachers' assignment to schools is managed by a central administration. One of the objectives of this reassignment process is to make sure that teachers obtain an assignment which they weakly prefer to their current position. The Deferred-Acceptance mechanism (DA) proposed by Gale and Shapley (1962) fails to satisfy this constraint. As a solution, a variation on this mechanism has been proposed in the literature and used in practice - as for the assignment of French teachers to schools. In this paper, we show that this mechanism yields assignments that can be improved in terms of both efficiency and "fairness". For each of the two efficiency notions considered in the literature (two-sided or one-sided), we identify the class of mechanisms which cannot be improved upon in terms of efficiency and fairness. Additionally, for two-sided efficiency, we show that a unique mechanism in the associated class is strategy-proof. For one-sided efficiency, no mechanism in the class is strategy-proof. Finally, using a rich dataset on teachers' applications to transfer in France, we empirically assess the extent of potential efficiency and fairness gains associated with the adoption of our mechanisms. These empirical results confirm the poor performance of (the variation on) the DA mechanism, and the significant improvement brought by our proposed mechanisms in terms of both efficiency and fairness.

*We are grateful to... and seminar participants at... for helpful comments.

†Paris School of Economics, France. Email: julien.combe@psemail.eu.

‡Paris School of Economics, France. Email: tercieux@pse.ens.fr.

§Paris School of Economics, France. Email: camille@cterrier.com.

1 Introduction

In many countries, the labor market for civil servants is highly regulated by a central administration. This is true for teachers whose assignment to public schools is often managed centrally by a public administration. This is the case in France but also in Mexico (Pereyra (2013)) Turkey (Dur and Kesten (2014)) or Uruguay (Vegas, Urquiola, and Cerdàn-Infantes (2006)). Given teachers' reported wishes and schools' preferences, an assignment is chosen by the administration via a matching mechanism. In France, each year, about 65,000 tenured teachers ask for an assignment.[1] In Turkey 8,850 positions were filled by new teachers in 2009 (Dur and Kesten (2014)). Teachers submit ranked lists of schools to a public administration and each school ranks teachers.[2] The criteria used to rank teachers are diverse, ranging from teachers' performance to a standardized test (in Turkey and Mexico) to teachers' experience or geographical distance to her/his partner (in France). In this paper, we focus on two standard requirements, namely efficiency and fairness, to assess performances of matching mechanisms. We theoretically identify mechanisms having the best performances in terms of these two criteria and empirically assess our findings.

The problem of assigning teachers to positions is quite similar to that of the assignment of students to schools. Yet, an important feature differs between these two problems. Contrary to students which are all initially unmatched, there are two types of teachers asking for an assignment: the newcomers who newly graduated and apply to be assigned for the first time; and teachers who already have a position and are willing to be re-assigned another position. In practice, tenured teachers have the right to keep their initial position if they wish, so that the administration has to offer these teachers positions that they weakly prefer to the school they are currently assigned to. In other words, the assignment

[1] Additionally, every year new teachers ask for their first assignment. About 10,000 non-tenured teachers ask for a first position every year.

[2] In France, three main criteria determine a teacher's ranking at schools: a teacher's total seniority, whether the teacher is currently assigned a disadvantaged/violent school, and whether a teacher demands for a spousal reunification. The latter criteria applies when a teacher asks for a transfer in order to live closer to his/her partner. In doing so, he/she obtains additional points proportional to the time they were separated.

of teachers must be individually rational. Distinguishing teachers who have a position from those who have no initial position has important consequences. While stability has emerged as a key feature to the success of many matching markets, in the realistic context where there are tenured teachers willing to be reassigned, all stable matchings may violate the individual rationality constraint.[3] Recall that a matching is stable if there is no blocking pair, that is, a pair of agents (say a teacher and a school) who prefer to match with each other rather than accepting the current matching. The most prominent stable mechanism is the Deferred Acceptance mechanism (DA, for short), in particular, because it gives incentives to teachers to truthfully report their preferences (i.e., it is strategy-proof). [4] But since it is stable, in our context where some teachers are initially assigned a school, DA fails to be individually rational.

Thus, an outstanding problem faced by the designer is to find, in a teacher assignment context, mechanisms which give better schools to teachers than their initial one and which keep "as much as possible" the nice properties of DA. As a solution, a variation of DA which ensures individual rationality has been proposed in the literature. It consists in artificially modifying schools' preferences so that all teachers initially assigned to a school are ranked at the top of that school's ranking. Once done, DA is run. By construction, this matching mechanism is individually rational. However, it has important flaws. We show that the assignment obtained can be improved simultaneously in two dimensions: efficiency in a very strong sense can be improved (i.e., in terms of both teachers' and schools' welfare) while at the same time the set of blocking pairs can be shrinked. Given this observation, one may naturally seek for mechanisms which do not have such flaws. The current paper characterizes these mechanisms, study their properties – in particular, in terms of incentives – and empirically assesses the theoretical findings using a rich data set on the assignment of teachers to public schools in France where the mechanism at use

[3]For cases with only first-year teachers (without an initially assigned position), the problem is formally equivalent to a two-sided matching problem as studied extensively since the seminal contribution by Gale and Shapley (1962).

[4]It is indeed, used in many real assignment problems. For instance in New York City for the assignment of students to high schools. A version of DA (NH4) is used in the assignment of on-campus housing at MIT, see Guillen and Kesten (2012).

is the modified version of DA mentioned above.

When characterizing alternative mechanisms, one has to think about the notion of efficiency to achieve. There are two ways to define efficiency in the literature. In the "school choice" approach (Abdulkadiroglu and Sonmez (2003)), schools' orderings over students are considered as priority orderings determined by laws and education policies, and are not necessarily seen as reflecting some welfare criterion of the designer. In the "college admission" approach (Gale and Shapley (1962)), schools' orderings are seen as true schools' preferences, so that both schools and students' welfare matters. Of course, schools are institutions and so having a designer who assumes that schools' "preferences" matter can only be justified if eventually the satisfaction of these preferences yield a better welfare to "real" agents. This is the way we see the status of schools' preferences in the college admission problem. More precisely, instead of considering students' and schools' welfare separately, an alternative way to think about the welfare of different stakeholders in a matching is to consider a policy maker who wants to maximize a social welfare function. Different elements would be weighted in this function, which eventually contribute to real agents' welfare. For instance students' obtained rank or their distance to home (students' preferences) could be part of the criteria, but also the social and ethnic diversity in schools (that often enters schools' rankings). Similarly, in our context of teachers' assignment, the way priorities of schools are defined is highly related to welfare considerations, even though priorities are defined by the law. Implicitely, the ranking given to teachers represent the criteria entering the social welfare function of the policy maker.[5] Thus, our interpretation of the efficiency criterion used in the college admission problem is directly related to this social welfare function: if one takes two different matchings such that the first one Pareto-dominates the second for both teachers and schools, the social welfare objective of the policy maker is better fulfilled with the former. Thus, our objective when characterizing

[5] As already mentioned, several criteria are used to determine schools' rankings over teachers. First teachers' experience is a criterion which is directly related to welfare considerations since it is well known that experienced teachers benefit to students (Rockoff (2004), Hanushek, Kain, and Rivkin (2005)). Other criteria are also used. Teachers have a higher priority if they ask a school which is closer to their spousal (spousal reunification criteria). Again, this can be directly related to teachers' welfare. If a school obtains a higher ranked teacher due to additional points given for the spousal reunification, the teacher will be matched to a school located closer to his partner which is welfare improving.

mechanisms is to avoid matchings that can be improved upon for both sides (we refer to them as two-sided maximal matchings) and hence cannot be improved in terms of fairness as well. We will also consider an efficiency notion in line with the school choice approach which – given our interpretation – corresponds to a welfare function with weights put mainly on rankings obtained by teachers. In that case, our objective is to characterize matching mechanisms that cannot be improved in terms of efficiency (on the teacher side) as well as in terms of fairness.

To characterize two-sided maximal matchings, we provide an algorithm, the Block Exchange (BE) algorithm. The idea is simple: if two teachers block with each others' schools, we allow these teachers to exchange their schools. Obviously, larger exchanges are possible involving many teachers. Implementing cycles in an appropriately built graph allows us to characterize all possible exchanges that can be made to obtain matchings that cannot be improved in terms of efficiency for both sides of the market. While the modified DA does not belong to this class, this class of mechanisms remain very large. Indeed, it is characterized by the collection of (two-sided) Pareto-efficient mechanisms that are individually rational on both sides of the market. However, we show that this large class of mechanisms contains a unique strategy-proof mechanism. We call this mechanism the Teacher-Optimal Block-Exchange (TO-BE). It can be described in a simple way using the well-known top-trading cycle mechanism (Shapley and Scarf (1974)) – TTC, for short. To describe specifically this mechanism, define the opportunity set of a teacher as the set of schools which rank him (weakly) above their initial assignments. Now, artificially modify each teacher's preferences so that each school outside his opportunity set is ranked as unacceptable. Then run TTC using these modified preferences. This mechanism turns out to be the unique strategy-proof mechanism in our class when the efficiency notion considered is two-sided.

Second, we characterize one-sided maximal matchings. If the gains of improving teachers' welfare are significant enough to overcome the possible losses in schools' welfare (which is a component of the policy maker objective), the social welfare objective could still be improved. Hence, our second approach focuses on matchings that cannot be improved in terms of teachers' welfare and stability (one-sided maximality). We provide an algo-

rithm called the One-Sided Block Exchange algorithm (1S-BE) that characterizes all the one-sided maximal matchings. Interestingly, while favoring teacher's welfare usually helps for incentives, we show that when considering one-sided efficiency (i.e., based only on teachers), the associated class of mechanisms contains no strategy-proof mechanism.

While our results identify classes of mechanisms that cannot be improved upon in terms of efficiency and fairness, they do not allow to quantify the possible performances across mechanisms in these classes neither do they allow to compare the gains of our proposed mechanisms compared to the modified DA. For instance, one could imagine a mechanism that cannot be improved in terms of efficiency and fairness but is inferior to DA in terms of utilitarian welfare and/or number of blocking pairs. Indeed, one can always construct (specific) examples where this type of phenomenon occurs. Therefore, in order to make progress, we add some structure to the environment. We assume that preferences and schools' rankings are drawn randomly from a rich class of distributions – covering a whole spectrum from complete correlation to total independence – and we study the statistical properties of our mechanisms when the number of agents gets large. With this additional structure, we are able to give a precise sense in which our mechanisms "perform quantitatively better" (i.e., in terms of utilitarian efficiency as well as number of blocking pairs) than the modified DA. We also identify the cost that the adoption of the unique strategy-proof mechanism could have in terms of utilitarian outcomes and number of blocks compared to a first best approach where one could pick any mechanism. Our arguments build on technique from random graph as in Lee (2014) or Che and Tercieux (2015a) and Che and Tercieux (2015b).

Finally, we empirically estimate the magnitude of gains and trade-offs in a real teachers' assignment problem. To do so, we use a rich data set on the assignment of teachers to public schools in France. Based on reported preferences by teachers (given that the mechanism at use is the modified version of DA mentioned above which is strategy-proof), we run counter-factuals to quantify the performances of our mechanisms. In the first part of the empirical analysis, we consider the class of mechanisms producing outcomes that cannot be improved when the welfare notion is two-sided. We confirm that the modified version of DA (DA* for short) can be simultaneously improved in terms of welfare (on both

sides of the market) and blocking pairs. Over the 49 markets (i.e., disciplines) where we ran DA*, 29 of them could be simultaneously improved in these two dimensions. Importantly, these 29 fields represent almost all the market we are analyzing since they contain 96.91% of teachers in our data set. In addition, the gains obtained by our alternative mechanisms are significant in the two dimensions. We show that, compared to DA*, the number of teachers moving from their initial assignment is more than doubled under our mechanisms. Additionally, under our mechanisms, the distribution of ranks of teachers (over schools they obtain) stochastically dominate that of DA*. Regarding stability, the number of teachers who are not blocking with any school increases by 36%. Finally, the percentage of schools having all of their positions improved doubles, going from 12.7% under DA* to 25.6% under our proposed mechanisms. These figures are essentially the same for the unique strategy-proof mechanism mentioned above, which makes it particularly appealing and a natural candidate to be implemented in practice.

The second part of the empirical analysis considers the mechanisms producing outcomes that cannot be improved when the welfare notion is only one-sided. As expected, relaxing the constraint allows for better performances in terms of teachers' welfare and number of blocking pairs, but this is done at the expense of schools' welfare. Compared to the mechanisms based on two-sided efficiency, the number of teachers who are not part of a blocking pair increases by 16% and the distribution of teachers' ranks stochastically dominates that of the mechanisms considered in the first part. However, 15.8% of the schools have at least one of their seats with a lower priority and 9.15% have a majority of seats with a lower priority.

The rest of the paper is organized as follows. We discuss the related literature below. Section 2 provides stylized example which motivates our analysis and introduces our basic definitions. Section 3 provides the theoretical analysis: it identifies the two classes of mechanisms previously described and studies the properties of the associated mechanism in particular in terms of incentives. Section 4 describes our data set on the the assignment of teachers to schools in France and reports the results of our counter-factuals. The last section concludes.

Related Literature

Our theoretical setup in this paper covers two standard models in matching theory. First, the *college admission problem* as defined by Gale and Shapley (1962). In this context, schools' have preferences that are taken into account for both welfare considerations as well as fairness/stability issues. Second, our model also embeds the *house allocation problem* as developed by Shapley and Scarf (1974). In this framework, individuals own a house and are willing to exchange among them their initial assignment. Hence, in this problem, only one side of the market has preferences. Among other things, one goal is to ensure that all individuals eventually get an assignment that they weakly prefer to their initial assignment. This problem is very similar to ours but, in our context, we do want to take into account the school side in a way which is similar to that of the college admission problem. While covering important applications, this "mixed "model has only been studied by a small number of authors. Guillen and Kesten (2012) is one of the exceptions and points out that the modified version of the DA mechanism is used for the allocation of on-campus housing at MIT. Compte and Jehiel (2008) and Pereyra (2013) provided results on the properties of this mechanism. They do point out that stability and individual rationality are not compatible. They propose a weakening of the notion of blocking pairs and show that the modified version of DA "maximizes stability"under their weakening. On the contrary, our work keeps the standard definition of blocking pairs and deals with maximal stability notions using the usual definition. More importantly, our theoretical and empirical results highlight the high cost that maximizing their stability notion can have in terms of efficiency and in terms of the traditional stability notion.

2 Basic Definitions and Motivation

Consider a problem where a finite set of teachers T has to be assigned to a finite set S of schools. For now, we restrict our attention to a one-to-one setting, i.e., an environment where each school has a single seat (see Section 4 for the treatment of the many-to-one case). Each teacher t has a strict preference relation $\succ_t$ over the set of schools and being unmatched (being unmatched is denoted by $\emptyset$). Similarly, each school s has a strict

preference relation $\succ_s$ over teachers and being unmatched. For any teacher t, we write $s \succeq_t s'$ if and only if $s \succ_t s'$ or $s = s'$. For any school s, we define $\succeq_s$ in a similar way. For simplicity, we assume that all teachers and schools prefer to be matched rather than being unmatched. A **matching** μ is a mapping from $T \cup S$ into $T \cup S \cup \{\emptyset\}$ such that (i) for each $t \in T$, $\mu(t) \in S \cup \{\emptyset\}$ and for each $s \in S$, $\mu(s) \in T \cup \{\emptyset\}$ and (ii) $\mu(t) = s$ iff $\mu(s) = t$. That is, a matching simply specifies the school where each teacher is assigned or if a teacher is unmatched. It also specifies the teachers assigned to each school, if any. We will also sometimes use the term assignment instead of matching. So far our environment is not different from the college admission problem (Gale and Shapley (1962)). However, in a teacher assignment problem, there is an additional component: teachers have an initial assignment. Let us denote the corresponding matching by μ_0. For now, we assume that $\mu_0(t) \neq \emptyset$ for each teacher t and $\mu_0(s) \neq \emptyset$ for each school s.[6] Hence, initially all teachers are assigned a school (there is no incoming flow of teachers) and there is no available seat at schools (there is no outgoing flow of teachers). We define a **teacher allocation problem** as a triplet $[T, S, \succ]$ where $\succ := (\succ_a)_{a \in S \cup T}$.

We will be interested in different efficiency and fairness criteria. Depending on whether we consider both teachers and schools as welfare-relevant entities or only the teacher side. First, we say that a matching μ is **two-sided individually rational** (2-IR) if for each teacher t, $\mu(t)$ is acceptable to t, i.e., $\mu(t) \succeq_t \mu_0(t)$ and, in addition, for each school s, $\mu(s)$ is acceptable to s, i.e., $\mu(s) \succeq_s \mu_0(s)$. Similarly, a matching is **one-sided individually rational** (1-IR) if each teacher finds his assignment acceptable. [7] We say that a matching μ 2-Pareto dominates (resp. 1-Pareto dominates) another matching μ' if all teachers and schools (resp. teachers) are weakly better-off – and some strictly better – under μ rather than under μ'. A matching is **two-sided Pareto-efficient** (2-PE) if there is no other

[6]This implies that μ_0 defines a bijection from T to S and so $|T| = |S|$.

[7]We assume strict preferences for simplicity. In the standard college admission problem, a school is acceptable if it is preferred to not being matched. The notation $\emptyset$ is usually used to denote the *null school* i.e. not being matched. Our model and definitions include the standard ones once we set $\mu_0(t) = \emptyset \; \forall t \in T$. Now, the statu-quo for a teacher is to stay at his initial school defined by μ_0 rather than being unmatched, so that the relevant comparison for acceptability is now with respect to the initial allocation. Our concepts and algorithms can be generalized to a setting where some teachers and/or schools are initially not matched, i.e. $\mu_0(a) = \emptyset$ for some $a \in T \cup S$.

matching which 2-Pareto dominates it. Similarly, we define **one-sided Pareto-efficient** (1-PE) matchings as assignments for which no alternative matching exist which 1-Pareto dominates it. We say that under matching μ, a teacher t has justified envy for teacher t' if t prefers the assignment of t', i.e., $\mu(t') =: s$, to his own assignment $\mu(t)$ and s prefers t to its assignment. Using the standard terminology from the literature, we say that (t, s) **blocks** matching μ. A matching μ is **stable** if there is no pair (t, s) blocking μ. We will sometimes say that a matching μ dominates another matching μ' in terms of stability if the set of blocking pairs of the μ is included in that of μ'.

Finally, a **matching mechanism** is a function φ which maps problems into matchings. We abuse notations and write $\varphi(\succ)$ for the matching obtained in the problem $[T, S, \succ]$. We will also note $\varphi_t(\succ)$ for the school that teacher t obtains under matching $\varphi(\succ)$. It is 2-IR/1-IR/1-PE/2-PE/stable if for each problem, it systematically selects a matching that is 2-IR/1-IR/1-PE/2-PE/stable.

One of the most classical matching mechanism is Deferred Acceptance (DA for short) which has been proposed by Gale and Shapley (1962). Since we will discuss a closely related mechanism, we recall its definition first.

- **Step 1.** Each teacher t applies to his most preferred school. Each school tentatively accepts its most preferred teacher among the offers it received and rejects all other offers.

In general,

- **Step $\mathbf{k \geq 1}$.** Each teacher t who was rejected at step $k - 1$ applies to his most preferred school among those to which he has not applied yet. Each school tentatively accepts its most favorite teacher among the new offers of the current step and the applicant tentatively selected from the previous step (if any) and it rejects all other offers.

The following proposition is well-known.

Proposition 1 (Gale and Shapley (1962)) *DA is a stable and 2-PE mechanism.*

While DA is stable and 2-PE, it fails to be 1-IR (and so 2-IR). As it turns out, this is unavoidable: in general, there is a conflict between individual rationality and stability. The basic intuition is that imposing 1-IR to a mechanism yields situations where some teacher t may be able to keep his initial assignment $\mu_0(t) =: s$ while school s may perfectly prefer other teachers to t. These other teachers may rank s at the top of the their preference relation and hence block with school s. We summarize this discussion in the following observation.[8]

Proposition 2 *There is no mechanism which is both 1-IR and stable. Hence, DA is not 1-IR.*

So there is a fundamental trade-off between 1-IR and stability and one may want to find a mechanism which restores individual rationality while keeping "as much as possible" the other nice properties of DA such as its stability and its 2-Pareto efficiency. In order to do so, one approach – followed by the literature, e.g., (see for instance Pereyra (2013) or Compte and Jehiel (2008)) and used in practice – consists in modifying artificially the schools' preferences so that each teacher t is ranked at the top of the (modified) ranking of the school he is initially assigned to, namely, $\mu_0(t)$. Other than this change, the schools' preference relations remain unchanged.[9] Once done, one runs DA as defined above using schools' modified preferences. We note the corresponding mechanism DA*. By construction, this is a 1-IR mechanism. This mechanism is used in practice in several situations. For instance, it is used for the assignment of on-campus housing at MIT (Guillen and Kesten (2012)). As described in Section 4, it is also used in France for the assignment of teachers to schools. Our empirical assessment of the paper's result will be based on this application.

By Proposition 2, we know that this mechanism is not stable. But is there a sense in which the violation of stability is minimal? What about in terms of efficiency: Is DA* 2-PE? And, if the answers to those questions are negative, can we find ways to improve

[8] This is already pointed out in Compte and Jehiel (2008) and Pereyra (2013)

[9] Formally, for each school s, a new preference relation $\succ'_s$ is defined so that $\mu_0(s) \succ'_s t'$ for each $t' \neq \mu_0(s)$ and for each t, t' distinct from the school's initial assigment $\mu_0(s)$, we have $t \succ'_s t'$ if and only if $t \succ_s t'$.

upon this? The following example will illustrate one important drawback of DA* on which we will come back both in our theoretical analysis as well as in our empirical assessment.

Example 1 *We consider a simple environment with n teachers and n schools. Let us assume that some teacher t^* is initially assigned to school s^* (i.e., $\mu_0(t^*) = s^*$) and is ranked first by all schools. In addition, school s^* is ranked at the bottom of each teacher's preference relation – including t^*, hence t^* is willing to move. We claim that under these assumptions no teacher will move from his initial assignment if we use DA* to assign teachers. To see this, note first that t^* does not move from his initial assignment. Indeed, because DA* is 1-IR, if t^* were to move then some teacher t would have to take the seat at school s^* but since s^* is the worst school for every teacher, this assignment would violate the individual rationality condition for teacher t, a contradiction. Note that this implies that under the DA* algorithm, t^* applies to every school s (but is eventually rejected) Now, to see that no teacher other than t^* moves, assume on the contrary that $t \neq t^*$ is assigned a school $s \neq \mu_0(t)$. As we already mentioned, at some step of the DA* algorithm, t^* applies to s. Since t^* is ranked above t in the preference relation of school s (recall that $s \neq \mu_0(t)$), t cannot eventually be matched to school s, a contradiction.*

Thus, to recap, under our assumptions, no teacher moves from his initial assignment. Since the initial assignment can perform very poorly in terms of basic criteria like stability or 2-Pareto efficiency, we can easily imagine the existence of alternative matchings which would make both teachers and schools better-off and hence shrink the set of blocking pairs as well.

The driving force in this example is the existence of a teacher ranked at the top of each school's ranking and who is initially assigned the worst school. This is of course a stylized example and one can easily imagine less extreme examples where a similar phenomenon would occur. The basic idea is that for DA* to perform poorly it is enough to have one teacher (a single one is enough!) being assigned an unpopular school and who himself has a fairly high ranking for a relatively large fraction of the schools. Our theoretical analysis as well as our empirical assessment will give a sense in which the described phenomenon is far from being a peculiarity. One may also argue that this effect is mainly driven by our assumption that every teacher is initially matched to a school and there are no empty

seats which may sound unrealistic in the context of teacher assignment. However, the underlying idea of avoiding possible movements in using DA* can still be found in such more realistic setting. The intuition relies on the congestion that can create correlation between teachers' preferences. If most of them tend to all agree on some top schools and even if top schools tend to be initially matched with top teachers, then these schools will have a huge set of applicants, some of them coming from other top schools and a large part coming from non top schools. Top schools' empty seats will be occupied by good applicants but still not all relatively good teachers coming from non top schools can be matched to these top schools due the IR constraint and the large number of applicants. So these teachers will create, as before, the same type of rejection effect under DA* for relatively bad teachers initially matched to the top schools and who would like to transfer to another top school. Even if the effect is soften, there is still a significant scope for improvement when it happens. We confirm this intuition in performing simulations in a more realistic environment with empty seats, non matched teachers, correlation in preferences and in the initial allocation. Details of the procedure and of the results can be found in Appendix A. Simulations seem to match the above intuition: correlation creates congestion in the top schools leading to a significant proportion of teachers not moving under DA*: for some parameters for instance, 62.5% of the initially matched teachers stay at their initial position under DA* and, over all iterations, we could improve upon these matchings for both sides of the market leading to 12% of additional teachers moving. When preferences are not extremely correlated, some exchanges remain possible between teachers not moving in the top schools so that DA* is not two-sided maximal. For almost completely independent or correlated preferences, almost no improvements for both sides of the market are possible.

The above example identifies a weakness of DA*: it can be improved both in terms of efficiency (on both side) as well as in terms of set of blocking pairs (i.e., we can shrink its set of blocking pairs). So we are interested in mechanisms/matchings which do not have this type of drawbacks. We also want to keep the elementary property that our mechanism/matching improves on the initial assignment. This suggests the following definitions.

Definition 1 *A matching μ is* ***two-sided maximal*** *if μ is 2-IR and there is no other*

matching μ' such that (1) all teachers and schools are weakly better-off and some strictly (2) the set of blocking pairs under μ' is a subset of that under μ.

This notion considers both schools and teachers as welfare-relevant entities. As we already argued, one may also ignore the school side. In such a case, we get the following natural counter-part.

Definition 2 *A matching is* ***one-sided maximal*** *if μ is 1-IR and there is no other matching μ' such that (1) all teachers are weakly better-off and some strictly (2) the set of blocking pairs under μ' is a subset of that under μ.*

Consistently with our previous notions, we say that a mechanism is two-sided (resp. one-sided) maximal if it systematically select a two-sided (resp. one-sided) maximal matching.

Let us note that if there is a matching μ' under which all teachers and schools are weakly better-off and some strictly than under a matching μ then the set of blocking pairs under μ' is a subset of that under μ. Thus, in the definition of two-sided maximality the requirement (2) can be dropped. This yields the following straightforward equivalent definition.

Proposition 3 *A matching μ is* ***two-sided maximal*** *if and only if μ is 2-IR and 2-PE.*

However, one can easily check that (2) in the definition of one-sided maximality cannot be dropped. Given Example 1 above, we have the following straightforward proposition.

Proposition 4 *DA* is not two-sided maximal and hence not one-sided maximal. Thus, DA* is not 2-PE.*

Given the above weaknesses of DA*, the obvious goal from now on is to identify the class of mechanisms characterizing two-sided as well as one-sided maximality and to study the properties of those mechanisms. The next section aims at doing so.

3 Theoretical Analysis

For each notion of maximality defined above (Definitions 1 and 2), the following two sections identify a class of mechanisms which characterize it. Once the characterization results proved, we analyze the properties of the mechanisms in that class. While the class of mechanism can be very large (as illustrated by Proposition 3), imposing standard additional conditions reduces drastically the set of candidate mechanisms. In particular, one striking outcome of this analysis is that, once the standard strategy-proofness notion is imposed, a unique two-sided maximal mechanism is shown to survive. In addition, while one may expect that giving more weight to teachers (as opposed to schools) as in one-sided maximal mechanisms may help in terms of incentive properties, another conceptually interesting outcome of our analysis is that no one-sided maximal mechanism is strategy-proof.

3.1 Two-sided maximality

In the next section, we define a class of mechanisms which characterizes the set of two-sided maximal mechanisms. As one may have expected, the mechanism will sequentially "clear" cycles of an appropriately constructed directed graph in the spirit of Gale's top-trading cycle, originally introduced in Shapley and Scarf (1974) and later studied by Abdulkadiroglu and Sonmez (2003).

3.1.1 The Block Exchange Algorithm

The basic idea behind the mechanisms we define is the following: starting from the initial assignment, if a teacher t blocks with the school initially assigned to t' and t' does also block with the school initially assigned to t, then we allow t and t' to "trade" their initial assignments. This is a pairwise exchange between t and t' but one may of course think of three way exchanges or even larger exchanges. Once such an exchange has been done, we obtain a new matching and we can look again at possible trades. More precisely, our class of mechanisms is induced by the following algorithm, named the Block Exchange (BE, for short):

- **Step** 0 : set $\mu(0) := \mu_0$.

- **Step** $k \geq 1$: Given $\mu(k-1)$, let the teachers and their assignments stand for the vertices of a directed graph where for each pair of nodes (t, s) and (t', s'), there is an edge $(t, s) \longrightarrow (t', s')$ if and only if teacher t blocks $\mu(k-1)$ with school s'. If there is no cycle, then return $\mu(k-1)$ as the outcome of the algorithm. Otherwise, select a cycle in this directed graph. For each edge $(t, s) \longrightarrow (t', s')$ in the cycle, assign teacher t to school s'. Let $\mu(k)$ be the matching so obtained. Go to step $k+1$.

It is easy to check that this algorithm converges in (finite and) polynomial time.[10] In the above description, we leave it open how the algorithm should select the cycle of the directed graph. Therefore, one may think of the above description as defining a class of mechanisms where a mechanism is determined only after we fully specify how to act when confronted with multiple cycles. One can imagine these selections to be random or dependent on earlier selections. In general, for each profile of preferences for teachers and schools, $\succ$, a possible outcome of BE is a matching that can be obtained by using an appropriate selection of cycles in the above procedure. Hence, we consider the following correspondence $BE : \succ \rightrightarrows \mu$ where $BE(\succ)$ stands for the set of all possible outcomes of BE. A selection of the BE Algorithm is a mapping $\varphi : \succ \mapsto \mu$ s.t. $\varphi(\succ) \in BE(\succ)$. Obviously, each selection φ of BE defines a mechanism.

As already mentioned, our class of mechanisms shares some similarities with Gale's top-trading cycle, however. There are two important differences. The first one, the most minor, is that a teacher in a node can point to several nodes and so implicitly to several schools. This is why contrary to top-trading cycle we do have an issue of selection of cycles and why our algorithm does not define a unique mechanism. However, as we will see in the next result, this is necessary for our characterization. Second, and certainly more importantly, our algorithm takes into account the welfare on both sides of the market.

[10]To see that this algorithm converges in a finite number of steps, observe that whenever we carry out a cycle, at least one teacher is strictly better-off. Hence, in the worst case one needs $(n-1)n$ steps for this algorithm to end. Since finding a cycle in a directed graph can be solved in polynomial time, the algorithm converges in polynomial-time.

Indeed, a teacher in a node (t, s) can point to a school in (t', s') only if s' agrees (i.e., s' prefers t to its assignment t'). This is what ensures, contrary to top-trading cycle, that each time we carry out a cycle, *both* teachers and schools become better-off. This has the nice implication that each time a cycle is cleared, the set of blocking pairs shrinks.

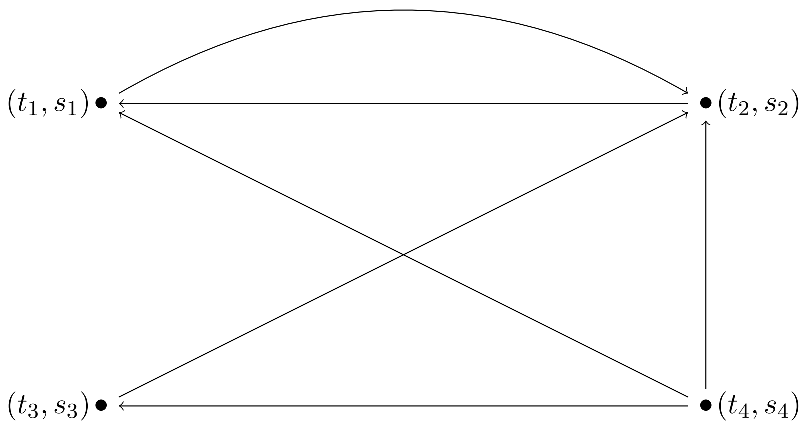
The BE algorithm starts from the initial assignment and then improves on it in terms of welfare of teachers and schools. More generally, one could start from any matching, obtained by running another mechanism φ. Doing so will guarantee the (modified) BE algorithm to select a matching which dominates that of φ both in terms of welfare of teachers and schools. This modification of the BE algorithm which takes the "composition" of BE and φ will be denoted by BE$\circ\varphi$. Given our starting point that DA* performs poorly in terms of welfare of teachers and schools, we will be particularly interested in BE$\circ$DA*.

The next example illustrates how the BE algorithm works.

Example 2 *There are 4 teachers $t_1, \ldots, t_4$ and 4 schools $s_1, \ldots, s_4$ with one seat each. The initial matching μ_0 is such that for $k = 1, \ldots, 4$, $\mu_0(t_k) = s_k$. Preferences are the following:*

$$\begin{array}{lllll} \succ_{t_1}: & s_2 & s_3 & s_1 & s_4 \\ \succ_{t_2}: & s_3 & s_1 & s_2 & s_4 \\ \succ_{t_3}: & s_1 & s_2 & s_3 & s_4 \\ \succ_{t_4}: & s_1 & s_2 & s_3 & s_4 \end{array} \qquad \begin{array}{lllll} \succ_{s_1}: & t_4 & t_2 & t_1 & t_3 \\ \succ_{s_2}: & t_4 & t_1 & t_3 & t_2 \\ \succ_{s_3}: & t_4 & t_3 & t_2 & t_1 \\ \succ_{s_4}: & t_4 & t_1 & t_2 & t_3 \end{array}$$

This example has a similar feature as Example 1: t_4 is the best teacher and is matched to the worst school. So we know that in that case: DA* coincides with the initial assignment. We have six blocking pairs: (t_1, s_2), (t_2, s_1), (t_3, s_2) and (t_4, s_k) for $k = 1, 2, 3$. The graph of BE is then the following:



The only cycle in this graph is $(t_1, s_1) \leftrightarrows (t_2, s_2)$ *and it can be checked that once implemented, there are no cycles left in the new matching so that the matching of BE is given by:*

$$BE = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 \\ s_2 & s_1 & s_3 & s_4 \end{pmatrix}$$

There are now 4 blocking pairs: (t_3, s_2) *and* (t_4, s_k) *for* $k = 1, 2, 3$ *and teacher* t_1 *and* t_2 *are both better-off.*

We now move to our characterization result.

Theorem 1 *Fix a preference profile. The set of possible outcomes of the BE algorithm coincides with the set of two-sided maximal matchings.*

Before proving the above statement we prove the following simple lemma.

Lemma 1 *Assume that* μ' *2-Pareto-dominates* μ*. Starting from* $\mu(0) = \mu$*, there is a collection of disjoint cycles in the directed graph associated with the BE algorithm which, once carried out, yields to matching* μ'*.*

Proof. Consider the directed graph where teachers and their assignments under μ stand for the vertices and for each pair of nodes (t, s) and (t', s'), there is an edge $(t, s) \longrightarrow$

(t', s') if and only if teacher t is assigned to s' under μ'. By definition of matchings, this directed graph has at least one cycle and cycles are disjoints. Note that because μ' 2-Pareto dominates μ, in this graph, $(t, s) \longrightarrow (t', s')$ if and only if teacher t blocks μ with school s'. Hence, the graph we built is a subgraph of the directed graph associated with the BE algorithm starting from μ. By construction, we have a collection of disjoint cycles in this directed graph which, once carried out, yields to matching μ', as was to be shown. ■

We are now in a position to complete the proof of Theorem 1.

Proof of Theorem 1. If μ is an outcome of BE, then it must be two-sided maximal. Indeed, if it were not the case, then by the above lemma, there would exist a cycle in the directed graph associated with the BE algorithm starting from μ which contradicts our assumption that μ is an outcome of the BE algorithm. Now, if μ is two-sided maximal, it 2-Pareto-dominates the initial assignment μ_0. Hence, appealing again to the above lemma, there is a collection of disjoint cycles in the directed graph associated with the BE algorithm starting from μ_0 which, once carried out, yield the assignment μ. Clearly, once μ is achieved by the BE algorithm, there is no more cycle in the associated graph. ■

While this result provides a simple and computationally easy procedure to find two-sided maximal matchings, the class of mechanisms defined by this algorithm is huge. Indeed, appealing to Proposition 3, this corresponds to the whole class of mechanisms that are both 2-PE and 2-IR. As we will see, by imposing the standard requirement of strategy-proofness, a unique mechanism will remain. The next section will state and prove this result as well as identify this mechanism.

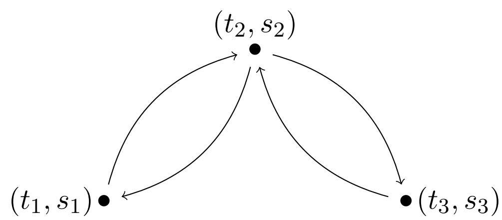
3.1.2 Incentives under Block Exchanges

First, recall that a mechanism φ is **strategy-proof** if for each profile of preferences $\succ$ and teacher t, $\varphi_t(\succ) \succeq_t \varphi_t(\succ'_t, \succ_{-t})$ for any possible report $\succ'_t$ of teacher t.[11] The following example shows that some selections of the BE algorithm are not strategy-proof.

Example 3 *Consider an environment with three teachers* $\{t_1, t_2, t_3\}$ *and three schools* $\{s_1, s_2, s_3\}$*. For each* $i = 1, 2, 3$*, we assume that teacher* t_i *is initially assigned to school*

[11] Using standard notations, $\succ_{-t}$ stands for the vector of preference relations $(\succ_{t'})_{t' \neq t}$.

s_i. Teacher t_1's most favorite school is s_2 while he ranks his initial school s_1 in second position. Teacher t_2 ranks s_1 first and then s_3. Teacher t_3 ranks s_2 first and then his initial assignment s_3. Finally, we assume that each teach is ranked in last position by the school he is initially assigned to. We obtain the following graph for the BE algorithm.



There are two possible cycles which overlap at (t_2, s_2). Consider a selection of the BE algorithm which picks cycle $(t_2, s_2) \leftrightarrows (t_3, s_3)$. In that case, the algorithm ends at the end of step 1 *and teacher t_2 is eventually matched to school s_3 his second most favorite school. However, if teacher t_2 lies and claims that he ranks s_3 below his initial assignment, the directed graph associated with the BE algorithm has a single cycle $(t_1, s_1) \leftrightarrows (t_2, s_2)$. In that case, the unique selection of the BE algorithm assigns t_2 to his most preferred school s_1. Hence, t_2 has a profitable deviation under the selection of the BE algorithm considered here.*

While this example is simple, one important objection for practical market design purposes is that the manipulation requires a fairly precise amount of information for teachers about the preferences in the market (i.e., for the other teachers as well as for schools). While this is true for many mechanisms, there is a sense in which – in some realistic instances – some selections of the BE (or associated) algorithm can be manipulated without the need of too much information on both preferences in the market as well as the details of the mechanism. A simple instance of this phenomenon can be illustrated for BE$\circ$DA*. Indeed, under this mechanism, a teacher who would initially be assigned a popular school which dislikes him can use the following strategy: reports his most preferred school sincerely and then ranks the school he is initially assigned to in second position (even though this may not match his true preferences). In case, the teacher does not get his first choice under DA*, he will certainly get his initial assignment under DA*.

Given that this school is popular and dislikes him, the teacher is likely to be part of a cycle involving his most favorite school when running the BE algorithm. Hence, at an intuitive level, this mechanism can be manipulated by teachers who may only have coarse information on preferences in the market.

In the following lines, we define a mechanism which is a selection of the BE algorithm and is strategy-proof. More surprisingly, we will prove further in the text that this is the unique selection satisfying this property. Before going through the definition of the mechanism, we need an additional piece of notation. Given a matching μ and a set of school $S' \subseteq S$, we let $\text{Opp}(t, \mu, S') := \{s \in S' \,|t \succeq_s \mu(s)\}$ be the opportunity set of teacher t within schools in S'. Note that for each teacher t, if $\mu_0(t) \in S'$, then $\text{Opp}(t, \mu_0, S') \neq \emptyset$ since $\mu_0(t) \in \text{Opp}(t, \mu_0, S')$.

- **Step** 0 : Set $\mu(0) = \mu_0$, $T(0) := T$ and $S(0) := S$.

- **Step** $k \geq 1$: Given $T(k-1)$ and $S(k-1)$, let the teachers in $T(k-1)$ and their assignments stand for the vertices of a directed graph where for each pair of nodes (t, s) and (t', s'), there is an edge $(t, s) \longrightarrow (t', s')$ if and only if teacher t ranks school s' first in his opportunity set $\text{Opp}(t, \mu(k-1), S(k-1)) = \text{Opp}(t, \mu_0, S(k-1))$. The directed graph so obtained is a directed graph with out-degree one[12] and, as such, has at least one cycle and cycles are pairwise disjoint. For each edge $(t, s) \longrightarrow (t', s')$ in a cycle, assign teacher t to school s'. Let $\mu(k)$ be the assignment obtained and $T(k)$ be the set of teachers who are not part of any cycle at the current step. If $T(k)$ is empty then return $\mu(k)$ as the outcome of the algorithm. Otherwise, go to step $k+1$.

As will become clear, our mechanism has a tight relationship with the top-trading cycle (TTC for short) mechanism. Recall that TTC works as the above mechanism except that

[12]Since, by construction, if t is not yet eliminated from the algorithm (i.e., he is in $T(k-1)$), so is the school to which t is initially assigned. Hence, $\mu_0(t) \in C(k-1)$. As we already noticed, this implies that $\text{Opp}(t, \mu_0, C(k-1))$ is non-empty. Now, because teachers have strict preferences, there is a unique most preferred school for t in $\text{Opp}(t, \mu_0, C(k-1))$.

the pointing behavior does not refer to the opportunity set: an edge $(t, s) \longrightarrow (t', s')$ is added if and only if teacher t ranks school s' first within the set of all remaining schools (i.e., at step k, those are the schools in $S(k-1)$). We will make use of the following straightforward equivalence result.

Lemma 2 *Fix a preference profile $\succ$. TO-BE($\succ$) is equal to TTC($\succ'$) where for each teacher t, the preference relation $\succ'_t$ ranks schools outside his opportunity set Opp(t, μ_0, S) below his initial assignment.*

From this simple lemma, we obtain the following proposition.

Theorem 2 *TO-BE is strategy-proof and is a selection of the BE algorithm.*

Proof. Given that an agent's report has no impact on his opportunity sets, Lemma 2 above (together with the well-known fact that TTC is strategy-proof) implies that TO-BE is strategy-proof. Now, we show that TO-BE is a selection of BE. Appealing to Theorem 1, it is enough to show that TO-BE is a two-sided maximal mechanism. If this were not to be the case, this would mean that for some preference profile $\succ$, starting from TO-BE($\succ$), there would be a cycle in the directed graph associated with the BE algorithm. It is easily checked that this cycle would still be there if preferences of teachers were to be modified so that any school outside the opportunity set of a teacher t (i.e., outside Opp(t, μ_0, S)) is ranked below his initial assignment. In this modified problem, by Lemma 2, TO-BE is equivalent to TTC. However, if we carry out the cycle starting from TO-BE we obtain a matching which Pareto-dominates for teachers TO-BE and hence TTC. This contradicts the well-know fact that TTC is 1-PE. ■

Say that a selection φ of the BE algorithm is teacher-optimal if there is no selection of BE which 1-Pareto-dominates φ. The following result justifies the terminology used so far: TO-BE is indeed teacher optimal.

Proposition 5 *Take any mechanism φ which is 2-IR. TO-BE is not 1-Pareto-dominated by φ.*

Proof. Proceed by contradiction and assume that TO-BE is 1-Pareto-dominated by φ at preference profile $\succ$. Since φ is 2-IR, $\varphi(t) \in$ Opp(t, μ_0, S). Hence, TO-BE is still

1-Pareto-dominated by φ at the modified preference profile where each teacher t ranks schools outside his opportunity set $\mathrm{Opp}(t, \mu_0, S)$ below his initial assignment. By Lemma 2, it implies that at the modified preference profile, TTC is 1-Pareto-dominated φ which is not possible given that TTC is 1-PE. ■

Corollary 1 *TO-BE is a teacher-optimal selection of BE.*

We now move to the most striking result of this section. Apart from TO-BE, no selection of the BE algorithm is strategy-proof.

Theorem 3 *TO-BE is the unique selection of the BE algorithm which is strategy-proof.*

Proof. The proof is relegated to Appendix B. ■

While the formal details of the argument are in the appendix, let us give a sktech of proof for this result.

As is well-known in a Shapley-Scarf economy (where schools are replaced by objects with no preferences but which are initially owned by the other side of the market) TTC is the unique element of the Core (Shapley and Scarf (1974) and Roth and Postlewaite (1977)). Because TO-BE is related to TTC, there is a sense in which it can be related to some Core notion. This notion is used in the course of the argument for Theorem 3. Define the two-sided Core notion as the set of matchings μ s.t. there is no (two-sided blocking) coalition $B \subseteq T$ for which there is a matching ν s.t. for each $t \in B$, $\nu(t)$ is a school to which a teacher in B is initally matched and for all $t \in B : \nu(t) \succeq_t \mu(t)$ and, for $s := \nu(t)$, $t \succeq_s \mu_0(s)$ with a strict equality for some teacher (or school). Given a profile of preferences, it is easily checked that a matching is in the two-sided Core if and only if it is in the (standard) Core when preferences are modified in such a way that each teacher t ranks schools outside his opportunity set $\mathrm{Opp}(t, \mu_0, S)$ below his initial assignment. Thus, appealing to the results mentioned above (i.e., Shapley and Scarf (1974) and Roth and Postlewaite (1977)), we conclude that the two-sided Core is a singleton and – given Lemma 2 – coincides with TO-BE.

Now, to give an intuition for Theorem 3, let us consider a selection φ of BE which is strategy-proof. Toward a contradiction, assume that φ and TO-BE differ at $\succ$. We first

prove a useful technical result: there exists a teacher t s.t. TO-BE$_t(\succ) \succ_t \varphi_t(\succ) \succ_t \mu_0(t)$. That there is a teacher who strictly prefers the assignment of TO-BE rather than that of φ is straightforward given that TO-BE is teacher-optimal (Proposition 1). The non-trivial part consists in showing that this very teacher also *strictly* prefers the assignment of φ rather than that of μ_0. If this was not the case, then among all teachers who strictly prefers TO-BE to φ, the assignment they would obtain with φ would concide with that of the initial assignment. Hence, if we denote B for the complement set of teachers, namely, those who weakly prefer the assignment given by φ rather than that given by TO-BE, we know that the assignment they obtain under φ corresponds to the initial assignment of some other teacher in B. Given that φ is 2-IR, this is very close to showing that B is a two-sided blocking coalition. To show that B is indeed a two-sided blocking coalition, we need to find a teacher in B who actually strictly prefers φ to TO-BE. Our argument shows that if this was not the case then this would contradict that φ is 2-PE (and so a selection of BE).

Now, given the above technical point, the proof proceeds as follows. Given the profile $\succ$, we consider modified preferences $\succ'_t$ for teachers which only rank as acceptable their school under TO-BE($\succ$). Given that this is the unique acceptable assignment for each teacher, the technical lemma implies that TO-BE($\succ'$) must be equal to $\varphi(\succ')$. We consider a sequence of unilateral deviations of teachers reporting $\succ_t$ instead of $\succ'_t$ which ultimately brings us back to $\succ$ and along which the equality between TO-BE and φ is maintained. To give an idea of why the equality is maintained along the sequence of unitalteral deviations, let us assume that, starting from $\succ'$, t reports $\succ_t$ instead of $\succ'_t$. If under $(\succ_t, \succ'_{-t})$, φ and TO-BE select different outcomes, then by the technical lemma again, we know that TO-BE$_t(\succ_t, \succ'_{-t}) \succ_t \varphi_t(\succ_t, \succ'_{-t}) \succ_t \mu_0(t)$.[13] By definition, TO-BE is not affected by t's deviation but then since TO-BE and φ coincide at $\succ'$, we have TO-BE$(\succ_t, \succ'_{-t}) = \varphi(\succ')$ which, by the previous argument, is strictly preferred to $\varphi_t(\succ_t, \succ'_{-t})$ at $\succ_t$. Thus, at $(\succ_t, \succ'_{-t})$, t can claim his preferences are $\succ'_t$ and get better-off, which contradicts the strategy-proofness of φ.

[13]Obviously, this condition cannot hold for teachers other than t' by construction of $\succ'$.

Hence, for an unilateral deviation of teacher t, φ and TO-BE must remain equal. Proceeding inductively in this way we can show that after a sequence of unilateral deviations from $\succ'_t$ to $\succ_t$ of each teacher, the equality between TO-BE and φ is maintained and, hence, TO-BE and φ coincide at $\succ$.

Before closing this section, we discuss the relationship to Ma (1994). Ma shows that in the Shapley-Scarf economy, the unique mechanism which is 1-IR, 1-PE and strategy-proof is TTC. Intuitively, our result applies to richer environments where schools have non-trivial preferences which are taken into account in the welfare. This suggests that our result is a generalization of Ma's. Indeed, to see this, note that in the specific situation where each school ranks its initial assignment at the bottom of its ranking, TO-BE and TTC coincide. In this context, 1-IR and 2-IR are obviously equivalent. In addition, since 1-PE implies 2-PE, we obtain that the class of mechanisms considered by Ma is a subset of the selections of the BE algorithm. Application of Theorem 3 to these selections yields Ma's result. While our argument builds upon that of Ma, there are a number of crucial differences. As already mentioned, even in the very specific environment where each school ranks its initial assignment at the bottom of its preference relation, the BE algorithm contains many other mechanisms which include in particular, all those that are 2-PE but not 1-PE and all 1-PE mechanisms that are "sensitive"to schools' preferences.[14] In addition, our result applies in general to settings where schools' preferences are arbitrary and so to many other types of mechanisms that are not even well-defined in Ma's environment. This requires to introduce non-trivial additional arguments to overcome the difficulties and prove Theorem 3.

3.2 One-sided maximality

We now turn to the characterization of one-sided maximality. As in the case of two-sided maximality, we introduce a class of mechanisms with possible outcomes spanning the whole set of one-sided maximal matchings.

[14]I.e., 1-PE mechanisms which select two different matchings for two different profiles of preferences where teachers' preferences remain unchanged.

While with two-sided maximality, the underlying criteria targeted by the designer are the welfare of teachers and schools as well as the set of blocking pairs, with one-sided maximality, the designer only targets the welfare of teachers and the set of blocking pairs. The basic idea behind the mechanism of this section is as follows: under the BE algorithm, two teachers can exchange their assignments iff they both block with the school initially assigned to the other teacher. However, one can imagine a pair of teachers t and t' who each desire the school of the other teacher – say s and s' respectively – and, while school s does not necessarily rank t' above t, it does rank first t' *among the individuals who desire* s.[15] If similarly, s' ranks first t among the individuals who desire s', then it is easily shown that an exchange between t and t' increases the welfare of teachers *and* shrinks the set of blocking pairs. Hence, based on a similar idea, we will weaken the definition of the pointing behavior in the directed graph defined in BE in such a way that – even though schools may become worse-off – both teachers' welfare increases and the set of blocking pair shrinks each time we carry out a cycle. The following algorithm – named one-sided BE (1S-BE for short) – accomplishes this weakening and Theorem 4 below gives a sense in which this is the best weakening one can hope for.

- **Step** 0 : set $\mu(0) := \mu_0$.
- **Step** $k \geq 1$: Given $\mu(k-1)$, let the teachers and their assignments stand for the vertices of a directed graph where for each pair of nodes (t, s) and (t', s'), there is an edge $(t, s) \longrightarrow (t', s')$ if and only if either (1) teacher t blocks with school s'; or (2) t desires s' and t is ranked first by s' among teachers who both desire s' and do not block with s'. If there is no cycle, then return $\mu(k-1)$ as the outcome of the algorithm. Otherwise, select a cycle in this directed graph. For each edge $(t, s) \longrightarrow (t', s')$ in the cycle, assign teacher t to school s'. Let $\mu(k)$ be the matching so obtained. Go to step $k+1$.

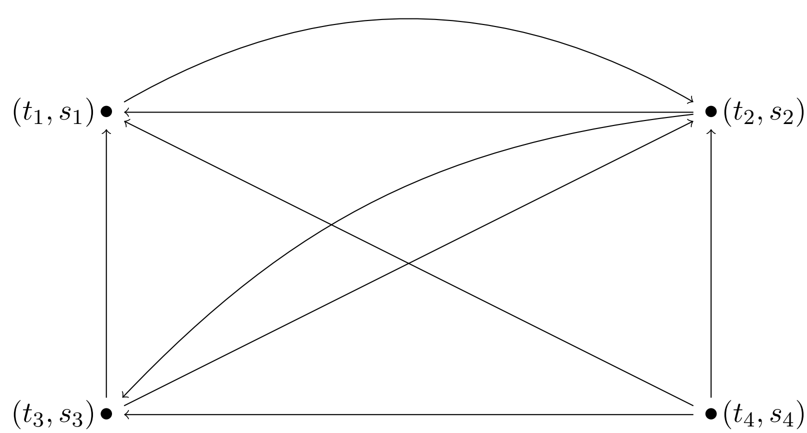
Here again, it is easy to check that this algorithm converges in (finite and) polynomial time. As for the BE algorithm, we leave it open how the algorithm should select the

[15] From now on, given a matching μ, we say that t desires s if $s \succ_t \mu(t)$.

cycle of the directed graph and so this algorithm defines a class of mechanisms. Each mechanism in this class is a selection of the correspondence from preference profiles to matchings corresponding to the whole set of possible outcomes that can be achieved by the 1S-BE algorithm.

By construction, starting from $\mu(k-1)$, the directed graph defined above is a supergraph of the directed graph that would have been built under the BE algorithm. Hence, there will be more cycles in our graph and more possibilities to improve teachers' welfare and to shrink the set of blocking pairs. This reflects the fact that we dropped the constraint that schools' welfare must increase along the algorithm and so more can be achieved in terms of teachers' welfare and set of blocking pairs. This is illustrated in the following example.

Example 4 *Take the same market as in Example 2. The graph of 1S-BE contains the edges of the graph of BE but it also has two new additional edges. Indeed,* t_1 *and* t_2 *both desire* s_3 *but do not block with it under* μ_0 *and* t_2 *is preferred to* t_1 *at* s_3 *so the node* (t_2, s_2) *can now point to* (t_3, s_3)*. Concerning* t_3*, he is the only one who desires* s_1 *and does not block with it so* (t_3, s_3) *can point to* (t_1, s_1)*. So the graph of 1S-BE is:*



Note that now, there are two additional cycle: $(t_1, s_1) \to (t_2, s_2) \to (t_3, s_3) \to (t_1, s_1)$ *and* $(t_1, s_1) \leftrightarrows (t_2, s_2)$*. Once having implemented the first one, it can be checked that there*

are no cycles left and so the matching given by 1S-BE is[16]

$$\begin{pmatrix} t_1 & t_2 & t_3 & t_4 \\ s_2 & s_3 & s_1 & s_4 \end{pmatrix}$$

Note that now, there are only three blocking pairs: (t_4, s_k) *for* $k = 1, 2, 3$.

Following the notions introduced for the BE algorithm, we will note 1S-BE$\circ\varphi$ for the "composition" of BE and of a mechanism φ. An outcome of such a (modified) 1S-BE algorithm selects matchings which dominate that of φ both in terms of welfare of teachers and set of blocking pairs (but not necessarily in terms of welfare of schools), i.e. all teachers are weakly better-off and the set of blocking pairs is a subset of that of φ. Again, in the sequel, we we will be particularly interested in starting the 1S-BE algorithm from the matching given by DA*, i.e., 1S-BE$\circ$DA*.

We now move to our characterization result. We note that while the argument in the proof of Theorem 1 is simple, the proof of the characterization result below is non-trivial.

Theorem 4 *Fix a preference profile. The set of possible outcomes of the 1S-BE algorithm coincides with the set of one-sided maximal matchings.*

Proof. The proof is relegated to Appendix C. ■

Assume that matching μ' dominates μ in terms of teachers' welfare and stability and consider the directed "exchange graph" where teachers and their assignments under μ stand for the vertices and for each pair of nodes (t, s) and (t', s'), there is an edge $(t, s) \longrightarrow (t', s')$ if and only if teacher t is assigned to s' under μ'. If μ' were to dominate μ in terms of teachers' *and schools'* welfare as well as in terms of stability, then, as argued in the proof of Lemma 1, each "cycle of exchange" in this graph is actually a cycle of the graph associated with the BE algorithm. This is core to the characterization result in Theorem 1. In the present case, where μ' dominates μ in terms of teachers' welfare as well as in terms of blocking pairs (but not necessarily in terms of schools' welfare), one may expect that these cycles of exchange would be cycles of the graph associated with the

[16] Note that even if one wanted to select one of the two other cycles, another cycle would lead to the same matching.

1S-BE algorithm. This turns out not to be the case and this is an important source of difficulty in the argument to prove Theorem 4. However, although cycles of exchange are not necessarily cycles of 1S-BE, we show that whenever there is a μ' which dominates μ in terms of teachers' welfare and stability, there must exist a cycle in the graph (which may not be a cycle of exchange) of 1S-BE starting from μ. With this existence, one direction of Theorem 4 can easily be proved. Indeed, given a matching μ obtained with the 1S-BE algorithm, if, toward a contradiction, it is not one-sided maximal, then, by definition, there must exist a matching μ' which 1-Pareto dominates μ and such that its set of blocking pairs is a subset of that of μ. But in that case, we know that there must exist a cycle in the graph associated with 1S-BE starting at μ which is a contradiction with the fact that μ is a matching obtained with the 1S-BE algorithm.

Here also, this result provides a computationally easy procedure to find one-sided maximal matchings. As for the BE algorithm, it is easy to construct selections of the 1S-BE algorithm which are not strategy-proof. In light of Theorem 3, an outstanding question naturally arises: is there any selection of the 1S-BE algorithm which is strategy-proof? While there is a unique selection of the BE algorithm which is strategy-proof, the next result provides a negative answer for the 1S-BE algorithm.

Theorem 5 *There is no selection of the 1S-BE algorithm which is strategy-proof.*

Proof. The proof is relegated to Appendix D. ■

This result points out an important difference between the classes of two-sided and one-sided maximal mechanisms. One can understand the difference as follows. Compared to the graph of BE, 1S-BE can have an edge $(t, s) \longrightarrow (t', s')$ if t desires s' and t is ranked first by s' among teachers who both desire s' and do not block with s'. Because of this condition, a teacher can modify the pointing behavior of others: indeed, if t is ranked first by s' among teachers who both desire s' and do not block with s', then teacher t can change the set of outgoing edges of other teachers depending on whether he claims that he desires s'. In the course of the argument for Theorem 5 we use this additional feature crucially. Indeed, we exhibit an instance under which, for each possible selection of cycles under the 1S-BE algorithm, one teacher can profitably misreport his preferences. Two

types of manipulations are used there: one is basic and consists in ranking as acceptable an unacceptable school in order to be able, once matched with it, to exchange it with a better one. However, for some selection of cycles, another manipulation is needed where a teacher ranks as unacceptable an acceptable school in order to expand the set of outgoing edges of other teachers. Again, this new type of manipulation is core to the argument in Theorem 5 and is not available under the BE algorithm.

Before closing this section, we note that the 1S-BE algorithm shares some similarities with the stable improvement cycle (SIC) algorithm defined by Erdil and Ergin (2008). Indeed, the 1S-BE could be seen as a generalization of the SIC algorithm. To discuss further the relationship, let us recall the definition of the SIC algorithm.

- **Step** 0 : set $\mu(0) := \mu_0$.
- **Step** $k \geq 1$: Given $\mu(k-1)$, let the teachers and their assignments stand for the vertices of a directed graph where for each pair of nodes (t, s) and (t', s'), there is an edge $(t, s) \longrightarrow (t', s')$ if and only if t desires s' and t is ranked first by s' among teachers who desire s'. If there is no cycle, then return $\mu(k-1)$ as the outcome of the algorithm. Otherwise, select a cycle in this directed graph. For each edge $(t, s) \longrightarrow (t', s')$ in the cycle, assign teacher t to school s'. Let $\mu(k)$ be the matching so obtained. Go to step $k+1$.

The SIC algorithm has been constructed to improve on stable outcomes whenever this outcome is not teacher-optimal as for instance is the case with the outcome of DA teacher proposing when schools have weak preferences. Now, note that when we start from a stable outcome SIC and 1S-BE are the same. Obviously, in our environment where schools have strict preferences, if we start from the outcome of DA – which here is the teacher-optimal stable assignment – the SIC and the 1S-BE do not have any cycle in their associated directed graph. More generally, if we were to weaken the assumption of strict preferences on the school side, the 1S-BE and the SIC algorithm – starting from DA – would yield to the same set of possible outcomes. However, our mechanism goes

much further in extending the properties of the SIC algorithm to cases where the starting assignment is arbitrary. To illustrate why this is true and why we cannot make use of the SIC algorithm for our purposes, consider one of our initial motivation which is to find ways to improve on the outcome of DA*. Both BE$\circ$DA* and 1S-BE$\circ$DA* succeed in doing so. However, the SIC algorithm (starting from the outcome of DA*) is of no help for this purpose. To see this, recall that under the SIC algorithm, $(t, s) \longrightarrow (t', s')$ iff t desires s' and t is ranked first by s' among teachers who desire s'. Since, by individual rationality of DA*, no teacher desires his initial assignment under the matching achieved by DA*, the pointing behavior – and so the directed graph – associated with SIC (starting from DA*) remains unchanged if we use the modified schools' preferences used to run DA* as opposed to the true schools' preferences. But under the modified preferences, by definition, DA* yields the teacher-optimal stable matching. Hence, there cannot be any cycle in the graph associated with SIC (again, starting from DA*).

3.3 Large Markets

Let us summarize our findings so far. We provided a stylized example where DA* performs poorly in terms of set of teachers moving from their initial position. This lack of movement naturally induces flaws for the assignments of DA*: one can improve on this algorithm both in terms of welfare of teachers and schools as well as set of blocking pairs. We have provided a whole class of mechanism – characterized by the BE algorithm – which does not suffer from such flaws. Hence, since lack of movement is an important source the weaknesses of DA*, one may ask the following question: *Is there more movement under all selections of the BE algorithm compared to DA*?*

While all selections of the BE algorithm are two-sided maximal, as pointed out in Proposition 3, two-sided maximality is a weak notion. One may naturally ask further how the different selections of the BE algorithm compare in terms of welfare of teachers and schools. In particular, we ask the following: *Based on standard welfare criteria, is there a best selection of BE?*

As we will see, there is a meaningful sense in which a best selection of BE exists in terms of welfare on both sides of the market. Since, one natural candidate mechanism –

the teacher-optimal BE – came out from our analysis on incentives, one may wonder how it compares with the best selection of BE. In a first best world where preferences would be known, one could use the best selection of BE. In a second best world where one adds the requirement of strategy-proofness, how do the welfare of teachers and schools is affected compared to the first best? Put in another way: *Is there a cost of strategy-proofness?*

This section will provide answers to these three questions. In order to make progress, we need to put more structure. Let us assume that we have K tiers for the schools. More precisely, there is a partition $\{S_k\}_{k=1}^{K}$ of S such that the utility of teacher t for school $s \in S_k$ $(k = 1, ..., K)$ is given by:

$$U_t(s) = u_k + \xi_{ts}$$

where $\xi_{ts} \sim U_{[0,1]}$. We assume that $u_1 > u_2 > ... > u_K$. For each $k = 1, .., K$, we denote x_k for the fraction of schools having common value u_k and further assume that $x_k > 0$. As for schools' preferences, we assume that

$$V_s(t) = \eta_{ts}$$

where $\eta_{ts} \sim U_{[0,1]}$. The additive separability structure of our utilities as well as the specific uniform distribution at use are not essential to our argument.[17]

Finally, the initial assignment μ_0 is selected at random among all possible $n!$ matchings where $n := |T| = |S|$. A random environment is hence characterized by the number of tiers, their size as well as common values $[K, \{x_k\}_{k=1}^{K}, \{u_k\}_{k=1}^{K}]$. The maximum normalized sum of teachers' payoffs that can be achieved in this society is $\bar{U}_T := \sum_{k=1}^{K} x_k (u_k + 1)$ which is attained if all teachers are matched to schools with which they enjoy the highest possible idiosyncratic payoff. The maximum normalized sum of schools' payoffs that can be achieved in this society is $\bar{U}_S := 1$ which is attained if all schools are matched to teachers with which they enjoy the highest possible idiosyncratic payoff. Clearly, in our environment where preferences are drawn randomly, a mechanism can be seen as a random

[17] We essentially need that the utilities are continuous and increasing in both components and that the distribution of the idiosyncratic shocks have full support in a compact interval in $\mathbb{R}$.

variable. In the sequel, we let $\varphi(t)$ be the random assignment that teacher t obtains under mechanism φ.

In general our mechanisms will fail to achieve the maximum sum of utilities on either side. However, a meaningful question is how often does this phenomenon occurs when the market gets large. The following concepts will help answering the question. We say that a mechanism φ **asymptotically maximizes movement** if, for any random environment,

$$\frac{|\{t \in T \,|\varphi(t) \neq \mu_0(t)\}|}{|T|} \xrightarrow{p} 1.$$

A mechanism φ is **asymptotically teacher-efficient** if, for any random environment,

$$\frac{1}{|T|}\sum_{t\in T} U_t(\varphi(t)) \xrightarrow{p} \bar{U}_T.$$

Similarly, φ is **asymptotically school-efficient** if, for any random environment,

$$\frac{1}{|S|}\sum_{s\in S} V_s(\varphi(s)) \xrightarrow{p} \bar{U}_S.$$

Finally, φ is **asymptotically stable** if, for any random environment, for any $\varepsilon > 0$,

$$\frac{|\{(t,s) \in T \times S \,|U_t(s) > U(\varphi(t)) + \varepsilon \text{ and } V_s(t) > V(\varphi(t)) + \varepsilon\}|}{|T \times S|} \xrightarrow{p} 0.$$

The next three results provide some answers to the three questions stated at the beginning of the current section. The proofs of these results are relegated in Appendix E.

Theorem 6 *DA^* does not maximize movement and hence it is neither asymptotically teacher-efficient, nor asymptotically school-efficient, nor asymptotically stable.*

The basic idea behind the above theorem is very close to the underlying argument in Example 1. Indeed, consider a random environment with two tiers of schools (i.e., $K = 2$) and where the second tier corresponds to "bad" schools (while the first corresponds to say "decent" schools). Formally, we assume that $u_1 > u_2 + 1$ so that irrespective of the idiosyncratic shocks, a school in tier 1 is always preferred to a school in tier 2. The intuition for the result is as follows. Fix any teacher t initially assigned to a school in the first tier. With non-vanishing probability, if t applies to a school in tier 1 other than his

initial assignment, some teacher in the second tier will be preferred by that school. Hence, teacher t will be kicked out by that teacher. This simple argument implies that – among teachers initially assigned to schools in tier 1 – the expected fraction of teachers staying at their initial assignment is bounded away from 0.

More specifically, for each $k = 1, 2$, let T_k stand for the set of teachers who are initially assigned a school in S_k. Consider any teacher $t \in T_1$. Let E_t be the event that for each school $s \in S_1$ there is a teacher $r \in T_2$ s.t. r is ranked above t (according to s's true preferences). Note that for a school s, the probability that t is ranked above each individual in T_2 is the probability that $\{t\} = \arg\max\{\eta_{ts}, \{\eta_{rs}\}_{r\in T_2}\}$. Since $\{\eta_{ts}, \{\eta_{rs}\}_{r\in T_2}\}$ is a collection of iid random variables, for each $r \in T_2$, by symmetry, the probability that the maximum is achieved by t must be the same as the probability that it is achieved by any $r \in T_2$. Hence, the probability of $\{t\} = \arg\max\{\eta_{ts}, \{\eta_{rs}\}_{r\in T_2}\}$ must be $\frac{1}{1+|T_2|}$. We can now easily compute the probability of E_t,

$$\begin{aligned}\Pr(E_t) &= \left(1 - \frac{1}{|T_2| + 1}\right)^{|S_1|} = \left(\left(1 - \frac{1}{|T_2| + 1}\right)^{|T_2|}\right)^{|S_1|/|T_2|} \\ &\to \left(\frac{1}{e}\right)^{x_1/x_2} \text{ as } n \to \infty.\end{aligned}$$

Note that,unsing the same logic as in Example 1, whenever E_t realizes, t cannot move from his initial assignment. Indeed, if t applies to some school s, this must be to a school in S_1. But, by construction, each teacher $t \in T_2$ applies to each school in S_1. In particular, the teacher in T_2 being ranked above t by school s applies to s, showing that eventually, t cannot be matched to s under DA*. Thus, the expected fraction of individuals in T_1 who do not move must be

$$\begin{aligned}\frac{1}{|T_1|}\mathbb{E}\left[\sum_{t\in T_1} \mathbf{1}_{\{t \text{ does not move}\}}\right] &= \frac{1}{|T_1|}|T_1|\,\mathbb{E}\left[\mathbf{1}_{\{t \text{ does not move}\}}\right] \\ &= \Pr\{t \text{ does not move}\} \\ &\geq \Pr(E_t).\end{aligned}$$

Thus, the $\liminf$ of the expected fraction of teachers not moving is bounded away from 0. Note that the lower bound computed here can be improved. Indeed, for t not to move,

one only needs that for each school $s \in S_1$ *that* t *finds acceptable*, there is a teacher $r \in T_2$ s.t. r is ranked above t (according to s's true preferences). In general, as shown below, our simulations suggest that a much large fraction of teachers are not moving and so the issue we point out is even more problematic.

Let us now think of the best possible outcome of the BE algorithm. While the way to implement this outcome may not be practical, we consider this as a benchmark and want to compare this to what can be achieved typically by mechanisms that can be implemented easily like TO-BE.

Theorem 7 *Each selection of BE asymptotically maximizes movement. There is a selection of BE which is asymptotically teacher-efficient, asymptotically school-efficient and asymptotically stable.*

The intution for the first part of the result is basic. Indeed, assume, toward a contradiction, that the set of teachers not moving under some selection of BE is "large". For each pair of teachers t and t' in that set, the probability t blocks with the initial assignment (and hence the assignment under the given selection) of t' and that, the other way around, t' blocks with the initial assignment of t (and so gain with the assignment under the given selection) is bounded away from 0. Hence, given our assumption tha the set of teachers not moving is large, with high probability, there will be such a pair of teachers. Put in another way, there will be a cycle in the graph associated with BE when starting from the assignment given by the selection, which contradicts the definition of a selection.

As for the other part of the theorem, within each tier, because preferences are drawn iid, one can build a random graph which, by results similar to Erdös-Rényi's existence of perfect matchings, ensures that with high probability, there is an assignment where both side of the market get high idiosyncratic payoffs.

Theorem 8 *TO-BE is asymptotically teacher-efficient. Under TO-BE, the expected payoff of a school is* $\frac{3}{4}$ *and so TO-BE is neither asymptotically school-efficient nor asymptotically stable.*

To understand why TO-BE is asymptotically teacher-efficient, an heuristic is as follows: assume it is not. This implies that for some tier, there is a "large set" of teachers who

are getting an idiosyncratic payoff bounded away from the upper bound. Now, for each pair of teachers t and t' in that set, intuitively, the probability that t blocks with the assignment of t' and that, the other way around, that t' blocks with the assignment of t is bounded away from 0. Hence, given our assumption tha the set of teachers not moving is large, intuitively, with high probability, there will be such a pair of teachers. Thus, here again, there will be a cycle in the graph associated with BE when starting from the assignment given by the selection, which contradicts the definition of a selection. While this is an intuitive way of describing the result, there is a difficulty here. Indeed, if we fix the set of teachers who are getting an idiosyncratic payoff bounded away from the upper bound, this has some implications on the distribution of preferences. Hence, there is a conditioning issue and the intuition given above does not take this into account. This raises an important technical difficulty and we go around this problem using random graph arguments in the spirit of those developed by Lee (2014) and more particularly Che and Tercieux (2015b).

From the above, we should expect from our data analysis several results. First, DA* should rarely be two-sided maximal, in particular, in markets with a large number of teachers. In addition, the BE algorithm and TO-BE should ensure more movement than DA* and perform better in terms of teachers' welfare. We will see that our data analysis largely confirm these findings. We should also expect TO-BE to perform less well than the BE algorithm: TO-BE may exhibit a loss in terms of schools' welfare and blocking pairs compared to the BE algorithm. In terms of schools' welfare and set of blocking pairs, it is not clear a priori how to compare TO-BE and DA*. Our data analysis will help discriminate further between these mechanisms.

4 Empirical Analysis

This section aims at assessing our theoretical findings by using a data set on the assignment of teachers to schools in France. We start by a brief description of the French procedure. We also provide a short presentation of the data set we collected. Finally, we run counterfactuals scenarios for our mechanisms and measure the extent of the improvements they may yield, both in terms of school and teacher's welfare as well as in terms of fairness.

4.1 The French teacher assignment system

A centralized procedure is used in France to assign public secondary school teachers to schools.[18] The central administration defines priorities over teachers using a point system, which takes into account three legal priorities: spousal reunification, disability, and having a position in a disadvantaged or violent school. Additional characteristics of teachers are also accounted for to compute the score: total seniority in teaching, seniority in the current school, time away from the spouse and/or children... This score determines schools' rankings or preferences – in this section we will use the terms priorities and preferences interchangeably). The point system is defined by the central administration, and well known by all teachers wishing to change school.[19]

The French territory is divided into 31 different regions which are called *académies*. We will refer to them as regions hereafter. Since 1999, the matching process has taken place in two successive phases:

Phase 1 (between regions assignment). New teachers and teachers who want to move to another region submit an ordered list of regions. A matching mechanism is used to match teachers to regions using priorities defined by the point system. This phase is managed by the central administration.

Phase 2 (within region assignment). For a given region, teachers matched to this region after the first phase and teachers who already have a position in the region but want to change school within their region report their preferences over the schools of their region. The same mechanism as in Phase 1 is used to do the match using priorities defined by a similar point system as in Phase 1.

The assignment process is decomposed in several markets - one for each subject taught - so that positions are specific to a subject and markets can be considered as being independent. In each market, the mechanism used in both phases is working as follows:

[18] For further details, the interested reader can read the description on the Matching in Practice website: http://www.matching-in-practice.eu/matching-practices-of-teachers-to-schools-france/

[19] An official list of criteria used to compute the point system is available on the government website http://cache.media.education.gouv.fr/file/42/84/6/annexeI-493_365846.pdf

first the priorities are artificially modified by moving teachers initially matched to regions (or schools) at the top of the priorities of their initial region (or school). Then a *School Proposing Deferred Acceptance* is run using the modified priorities and the reported preferences. Finally, *Stable Improvement Cycles* are executed as defined in Erdil and Ergin (2008) – see Section 3.2 for a description. Using Theorem 1 in Erdil and Ergin (2008), we know that this process will converge to the outcome of the *Student Proposing Deferred Acceptance* according to the modified priorities and reported preferences.[20] Hence, the mechanism used to assign French teachers to public schools is equivalent to DA* - as defined in Section 2.

In 2012, just over 26 000 teachers applied in Phase 1, and around 55 300 submitted a list to be assigned a new school within a region (i.e., in Phase 2). These figures include all newly tenured teachers, who have never been assigned a position, and tenured teachers who ask for a transfer. There are 107 different subjects taught, having different sizes in terms of teachers. Some are large like Sport (around 2500 teachers), Contemporary Literature (around 2000 teachers), Mathematics (around 2000 teachers), while some are smaller like Thermal Engineering (around 60 teachers) or Beauticians (around 15 teachers) with a wide range of subjects in between.

4.2 Data

We use several data sets related to the assignment of teachers in 2013. For both the first and the second phase of the assignment, these data sets contain four key information: (1) the reported preferences of teachers, (2) the priorities of regions/schools, (3) the initial assignment of each teacher (if any) and (4) the capacities of the regions/schools. As discussed in Appendix G, because of strategic issues, reported preferences during the

[20] In this paper we ignore the issue with weak priorities. However, priorities can be weak in the French system. Ties are broken using birth dates. Hence, to be more precise, tier are broken, then according to the modified priorities, school Proposing Deferred Acceptance. From this outcome Stable Improvement Cycles are run using the modified priorities *with no ties*. Hence, the outcome is equivalent the Student Proposing Deferred Acceptance which in turn may be Pareto-dominated by a Student-Optimal Stable mechanism. Defining our mechanisms in case of weak priorities is an easy exercise.

second phase of the assignment are not completely reliable – in particular, because of a binding constraint on the number of wishes that teachers can report – so that we restrict our analysis to the first phase: the assignment of teachers to the 31 regions. The sample of teachers used for the analysis takes into account two restrictions. First, the sample is restricted to the 49 subjects that contain more than 10 teachers. Second, in order to match our theoretical framework, all initially non matched teachers (newly tenured) and all empty seats in regions have been suppressed. Hence, the initial assignment corresponds to a market where each teacher is initially assigned a region and each seat of each region is initially assigned a teacher. The final sample contains 10 579 teachers corresponding to 49 subjects ranging from 6 to 1753 teachers.

4.3 Results

Before moving to the description of the results, we need to discuss briefly the generalization of our concepts to the many-to-one framework. The main challenge is to find a way to compare a school welfare across mechanisms in this new setting. Two-sided maximality, which is characterized by the BE mechanism, relies on the improvement of the welfare of both sides of the market, so that no school/region should be hurt. The following welfare definition will be used in the many-to-one framework : a school/region is **strongly better off** under matching μ' than under matching μ if at that school/region the teacher with the highest priority under μ' has a higher priority than the teacher with the highest priority under μ. Similarly, the teacher with the second highest priority under μ' has a higher priority than the one with the second highest priority under μ, and so on... This defines a strong improvement notion and we will discuss possible relaxations further in the text. We can then easily redefine our two-sided maximality concept using this welfare comparison for regions. In Appendix F, we provide the extensions of our mechanisms to the many-to-one framework. These generalizations are the ones used in the empirical analysis.

The following empirical analysis aims at testing our theoretical results. Therefore, we will focus on three main dimensions: teachers' welfare, regions' welfare and number of blocking pairs. Since BE and 1S-BE define a class of mechanism, we randomly select ten iterations in this class. For TO-BE, in a many-to-one environment, it is not true anymore

that there is only one outgoing edge from a given node. Indeed, teachers select their best school in their opportunity set. Since this school may have several seats, there may be several nodes involving the school so that some cycles may intersect. Hence, as for BE and 1S-BE, TO-BE defines a class of mechanism from which we randomly draw selections. We iterate the random selections for each mechanism ten times, and select the iteration of BE, TO-BE and 1S-BE which leads to (1) the maximum movement among teachers for results on teachers' and regions' welfare and (2) the minimum number of blocking pairs for results on stability.[21] The next section starts presenting results for the BE algorithm. Then the results for the 1S-BE are reported. .

4.3.1 Two-sided maximality: BE and TO-BE

How far is DA* from being two-sided maximal?

In the theoretical analysis, we have pointed out an important flaw of DA*: it can be improved in three main dimensions, namely, teachers' welfare, regions' welfare and fairness. The two-sided maximality concept captures this idea and is characterized by the BE algorithm. In practice, a simple way to test if DA* is two-sided maximal is to run the BE algorithm starting from the matching obtained by DA*, and then to observe if both matchings obtained with DA* and BE$\circ$DA* differ. If they differ, this means that some cycles exist in the graph associated with the BE algorithm starting from DA*, so that the latter is not two-sided maximal.

Fact 1 *DA* is not two-sided maximal in 22 subjects out of 49. These represent 96.91% of the whole set of teachers.*[22]

This first fact suggests that the theoretical phenomenon we highlight is not rare. Based on this observation, we now estimate how far DA* is from maximality in terms of the three criteria we are interested in. First, regarding teachers' welfare, two results reported in Table 1 are worth commenting. First, BE$\circ$DA* more than doubles the number of teachers

[21] Different choices of iterations do not qualitatively change our results.

[22] If we restrict our attention to the 19 subjects with more than 100 teachers, only two of them are not two sided maximal.

who obtain a new region compared to DA*: 563 teachers are moving from their initial allocation under DA* versus 1381 under BE∘DA*.[23] Second, the cumulative distribution of the rank obtained by teachers under BE∘DA* first order stochastically dominates this same distribution under DA*.

Table 1: Cumulative distribution of the number of teachers who obtain school rank k

Rank	Init	DA*	TO-BE	BE(Init)	BE(DA*)	1S-BE(Init)	1S-BE(DA*)
1	0	422	957	983	1068	1391	1378
2	7935	8084	8226	8283	8283	8402	8379
3	9125	9221	9289	9324	9335	9395	9405
4	9743	9797	9846	9873	9882	9928	9932
5	10038	10078	10113	10129	10135	10173	10171
6	10271	10297	10320	10323	10326	10354	10350
7	10366	10384	10402	10405	10407	10423	10420
8	10420	10433	10445	10445	10448	10460	10459
9	10461	10475	10481	10479	10486	10491	10495
>= 10	10579	10579	10579	10579	10579	10579	10579
Nb teachers moving		563	1028	1286	1381	1795	1756

† Notes: Data from the assignment process of French teachers to regions in 2013.

Second, regarding regions' welfare, we are interested in the percentage of positions, within a region, which are filled by a teacher having a higher priority than the teacher initially assigned to this position. Table 2 reports the cumulative distributions of this statistic for the different mechanisms we run. Again, we observe that the cumulative distribution under BE∘DA* first order stochastically dominates this same cumulative distribution under DA* [24].

[23] The limited share of teachers who move is mostly explained by the high proportion of teachers reporting short lists. Indeed, teachers rank on average 1.64 regions and 75 % of teachers ask only for one region (beyond their initial region). Combined with correlation in preferences, this structurally restricts the possibility of movement in the market. If teachers tend to ask for the same regions, and rank only a limited number of them, since the number of seats in this region is limited, many teachers have to stay at their initial allocation.

[24] Note that the bottom line of Table 2 corresponds to the percentage of regions being strongly better

Finally, we compare the performance of DA* and BE$\circ$DA* in terms of fairness. The first row of Table 3 reports that 2890 teachers are not blocking under DA* and 3977 under BE$\circ$DA*, which represents an increase of 37.6 % of the number of teachers which are not blocking with any region. Justified envy is reduced. [25]

All together, these results show that DA* fails to be two-sided maximal in a large number of cases and the scope of improvement seems to be very large. To tackle this issue, a first natural candidate could be to run the BE algorithm from the matching of DA*. However, as mentioned in our theoretical analysis, this mechanism is prone to easy manipulations. Alternatively, we focus our attention on both the BE algorithm which is run directly from the initial assignment (this is referred to as BE(Init) in our tables and graphs), and its strategy-proof selection : the TO-BE mechanism. In the next section, we evaluate the performance of these two mechanisms in terms of teachers' welfare, regions' welfare and number of blocking pairs.

Performance of BE and TO-BE

Before commenting the results, it is worth discussing briefly the relevance of comparing BE and TO-BE to DA*. We should keep in mind that an arbitrary outcome of the BE mechanism may not be comparable with that of DA* according to the two-sided maximality order: the set of blocking pairs may differ and the matchings may not be ranked according to the 2-Pareto domination order. However the comparison remains interesting for two reasons. First, we know by the above results that DA* is far from being two-sided maximal, so that BE and TO-BE can be expected to perform much better. Second, our theoretical results (Theorems 6, 7 and 8) suggest that BE and TO-BE

(w.r.t. the initial assignment), ie regions which have at least one position filled by a teacher having a higher priority, and no position filled by a teacher having a lower priority. BE$\circ$DA* doubles the number of regions which are strongly better off (w.r.t. the initial assignment) compared to DA*.

[25]Contrary to the number of teachers moving, the number of teachers being part of a blocking pair is quite high. This is intuitive since the number of teachers moving is low and many of them stay at their initial allocation creating possible envy. This can be seen as the cost of the individual rationality constraint.

perform better in large markets than DA*.

Table 2: Cumulative distribution of the percentage of regions having k % of their positions filled by a teacher with a higher priority (compared to the initial assignment).

% Positions	DA*	TO-BE	BE(Init)	BE(DA*)	1S-BE(Init)	1S-BE(DA*)
100%	1.71	4.48	5.27	6.25	7.44	8.43
90%	2.24	5.33	6.25	7.44	7.97	9.15
80%	3.95	7.31	8.69	9.87	9.35	10.86
70%	4.81	8.62	10.66	11.32	10.14	11.72
60%	6.39	10.86	12.90	13.56	11.65	13.23
50%	7.97	14.15	16.66	16.72	14.09	15.80
40%	8.82	15.73	18.76	18.30	14.75	16.52
30%	10.20	18.37	21.33	19.88	16.06	17.64
20%	11.06	21.20	23.24	21.92	16.79	18.56
10%	12.11	22.45	24.82	22.98	17.58	19.22
0%	12.71	23.44	25.61	23.70	17.84	19.62

† Note: Sample restricted to the strongly better regions, ie regions having at least one position filled by a teacher having a higher priority, and no position filled by a teacher having a lower priority. For each mechanism, we use the iteration leading to the highest movement. There are 31 regions in the sample. For each of the 49 subjects, we computed the number of stongly better regions and divided it by 49×31 to obtain the average percentage. Under DA* on average 7.97 % of the regions have at least 50 % of their positions filled by a teacher with a higher priority compared to the initial allocation, while under BE$\circ$ DA* this percentage reaches 16.72 %.

We first focus our attention on the performance of BE and TO-BE in terms of teachers' welfare. Both mechanisms significantly improve the number of teachers moving compared to DA* : 563 teachers obtain a new assignment under DA*, versus 1286 under BE and 1028 under TO-BE.

Fact 2 *The cumulative distribution of the rank obtained by teachers under BE and TO-BE first order stochastically dominates the cumulative distribution of DA*.*[26]

Note however, that there is no 2-Pareto domination between the matchings: some

[26] Distributions of BE and TO-BE are not stochastically ordered even if BE seems to perform generally better.

teachers may prefer their assignment under DA* to the one they obtain under BE or TO-BE.

Fact 3 *The three cumulative distributions of the number of teachers blocking with k regions under DA*, BE and TO-BE can be ranked stochastically: DA* is dominated by TO-BE which is dominated by BE. Justified envy decreases with each of these mechanisms.*

As suggested by our theoretical results, the domination of BE over TO-BE can be viewed as a cost of imposing strategy proofness.[27] Finally, comparing regions' welfare across mechanisms is of particular interest since our theoretical results highlight that BE and TO-BE are two-sided maximal, which is not the case of DA*. Thus, DA* should not perform well for regions compared to BE and TO-BE. The bottom line of Table 2 confirms that 12.7 % of the regions are strongly better off under DA*, while this percentage raises to 23.4% under TO-BE and 25.6% under BE.

Fact 4 *The cumulative distribution of the percentage of regions which have positions filled by a teacher they prefer can be stochastically ordered : the distribution of BE stochastically dominates that of TO-BE, which dominates that of DA* (Table 2).*

The lower performance of TO-BE compared to BE in terms of regions' welfare is in line with our theoretical predictions on the cost of the strategy proofness imposed by TO-BE. Finally, we know that DA* could even hurt some regions, contrary to BE and TO-BE. Results on this are presented in Appendix G.

All together, these results suggest that BE and TO-BE perform much better than DA* in terms of teachers' welfare, regions' welfare and fairness. The good performance of TO-BE is of particular interest due to its incentive properties. These results provide evidence that even though two-sided maximality is a strong requirement, our mechanisms can generate large improvements and distributions dominating those of DA*. [28] The next

[27] As discussed for teachers' welfare, it is worth noticing that the set of blocking pairs of each matching may differ. Some teachers may block with a region under BE or TO-BE but not under DA*.

[28] These results are all the more encouraging as they are obtained in a restrictive environment where teachers rank a very limited number of regions. Even better results could be expected in environments where agents have longer ranked lists.

section tests if we could further improve upon DA* by relaxing the constraint that no region should be hurt. We provide empirical evidence on the performance of 1S-BE, the one-sided maximal algorithm we have defined in Section 3.2.

Table 3: Cumulative distribution of the number of teachers with k blocking pairs

Nb Blocking Pairs	Init	DA*	TO-BE	BE(Init)	BE(DA*)	1S-BE(Init)	1S-BE(DA*)
0	2379	2890	3859	3948	3977	4580	4606
1	8899	9058	9262	9329	9338	9468	9505
2	9787	9880	9995	10046	10049	10112	10145
3	10167	10221	10299	10330	10345	10371	10383
4	10368	10400	10445	10460	10480	10489	10493
>= 5	10579	10579	10579	10579	10579	10579	10579

† Notes: Data from the assignment process of French teachers to regions in 2013.

4.3.2 One-sided maximality: 1S-BE

As done previously for BE and TO-BE, we first want to estimate how far DA* is from being one-sided maximal. To do so, we compare the matching under DA* to the one under 1S-BE that we run from DA*. For a large number of subjects, we have seen that DA* is not two-sided maximal so that it is not one-sided maximal either. Because the constraint on regions' welfare is relaxed under 1S-BE compared to BE, the improvements we have found for BE$\circ$DA* in terms of teachers' welfare and blocking pairs can be seen as a lower bound on improvements under 1S-BE$\circ$DA*. Indeed, Table 1 reports that 1S-BE$\circ$DA* multiplies by more than three the number of teachers moving compared to DA* and increases it by 27% compared to BE$\circ$DA*. This suggests that there is still significant possible improvements when considering one-sided maximality.

Fact 5 *DA* is not one sided maximal in 29 subjects. 98.54% of the teachers belong to these subjects*[29]*. For 7 subjects, DA* is two sided maximal but not one sided maximal (there are cycles in the graph associated with 1S-BE but not in that of BE).*

[29] If we restrict our attention to the subjects with more than 100 teachers, only one of them is not one sided maximal.

We now turn to the results on 1S-BE starting from the initial allocation (sometimes referred to as 1S-BE(Init) in our tables and graphs) to compare its performance with the other mechanisms. When running 1S-BE, we obtain similar results regarding teachers' welfare and fairness : the cumulative distributions of both the ranks obtained by teachers (Table 1) and the number of teachers blocking (Table 3) under 1S-BE stochastically dominates the distribution of all other algorithms mentioned previously: BE, TO-BE and DA*. This means that teachers' welfare and fairness can be further improved when we care less about regions' welfare[30].

Finally, regions' welfare is the key difference between two and one sided maximality. Even if improvements for teachers welfare and number of blocking pairs are large with 1S-BE, we know from the theoretical section, that this is done at the expense of the regions' welfare.

Fact 6 *Contrary the algorithms BE or TO-BE were regions' welfare cannot be hurt, under 1S-BE, 15.8 % of the regions have at least one position filled by a teacher having a lower priority than the teacher initially assigned to this position (Table 5 in Appendix G).*

Not only can regions be hurt under 1S-BE, but the welfare gains they experience (if any) are also smaller than under alternative algorithms. On average, 17.8 % of the regions are strongly better off under 1S-BE (Table 2), while this percentage raises to 25.6 % under BE and 23.4 % under TO-BE[31]. Additional results on regions' welfare are available in Appendix G. We take into account, within a region, the "net" number of positions being assignned teachers with higher and lower priorities.

Finally, as an alternative comparison, we also ran the Top Trading Cycles mechanism. Since it relaxes the constraint on the blocking pairs and mainly focuses on the welfare of teachers, TTC is an interesting algorithm to compare. Its performances are detailed in Appendix G.

[30] As explained before, some teachers may prefer their match under DA*. 195 teachers do so under 1S-BE, which is less than those under BE or TO-BE.

[31] Note however that the cumulative distribution of the percentage of regions having positions filled by teachers they prefer under 1S-BE is not stochastically dominated by this cumulative distribution under BE or TO-BE (Table 2).

References

Abdulkadiroglu, A., and T. Sonmez (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

Che, Y.-K., and O. Tercieux (2015a): "Efficiency and Stability in Large Matching Markets," Columbia University and PSE, Unpublished mimeo.

——— (2015b): "Payoff Equivalence of Efficient Mechanisms in Large Markets," Columbia University and PSE, Unpublished mimeo.

Compte, O., and P. Jehiel (2008): "Voluntary Participation and Re-assignment in Two-sided Matching," Paris School of Economics, Unpublished mimeo.

Dur, U., and O. Kesten (2014): "Pareto Optimal, Strategy-proof, and Non-bossy Matching Mechanisms," .

Erdil, A., and H. Ergin (2008): "What's the Matter with Tie-breaking? Improving Efficiency in School Choice," *American Economic Review*, 98(3), 669–689.

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

Guillen, P., and O. Kesten (2012): "Matching Markets with Mixed Ownership: The Case for a Real-life Assignment Mechanism," *International Economic Review*, 53(3), 1027–1046.

Haeringer, G., and F. Klijn (2009): "Constrained School Choice," *Journal of Economic Theory*, 144, 1921–1947.

Hanushek, E. A., J. F. Kain, and S. G. Rivkin (2005): "Teachers, schools, and academic achievement," 73, 417–458.

Lee, S. (2014): "Incentive Compatibility of Large Centralized Matching Markets," University of Pennsylvania, Unpublished mimeo.

Ma, J. (1994): "Strategy-Proofness and the Strict Core in a Market with Indivisibilities," *International Journal of Game Theory*, 23, 75–83.

Pereyra, J. S. (2013): "A Dynamic School Choice Model," *Games and economic behavior*, 80, 100–114.

Rockoff, J. E. (2004): "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," *American Economic Review*, 94, 247–252.

Roth, A. E., and A. Postlewaite (1977): "Weak versus Strong Domination in a Market with Indivisible Goods," *Journal of Mathematical Economics*, 4(2), 131–137.

Shapley, L., and H. Scarf (1974): "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1, 22–37.

Vegas, E., M. Urquiola, and P. Cerdàn-Infantes (2006): "Teacher Assignment, Mobility and their Impact on Equity and Quality of Education in Uruguay," Unpublished mimeo.

A Simulations

We used a more realistic framework in the simulations to check robustness of our result concerning DA*. We so have a market with 600 teachers and 30 schools with 20 seats each. The utility of school s for teacher t is given by:

$$U_t(s) = u_s + \eta_{st}$$

where u_s is the common shock of school s picked uniformly over the interval $[0, a]$ and η_{st} is the idiosyncratic shock picked uniformly over the interval $[0, 1]$. In changing the parameter a we can change the correlation in the preferences: if $a = 0$ there is no correlation and if $a = \infty$ preferences are the same with probability 1. For the schools priorities, we have the same specification except that both the common shocks (v_t) and the idiosyncratic shocks (ξ_{st}) are uniformly taken in the interval $[0, 1]$. So we would like to represent some realistic features of the teachers' assignment: each year almost the same number of teachers leave and enter the market so that there is a significant proportion of empty seats. Preferences tend to be correlated over the schools. Teachers with more priorities also tend to be initially matched to the preferred schools. In order to capture these features, we set the initial allocation as follows: we match the 10 best teachers according to their common shock in school priorities to the school having the highest common shock in teachers preferences. And we continue in this way: we match the 10 next best teachers to the school with the second highest common shock and so on. Note that in doing so, each school has 10 seats initially matched and 10 empty seats and among the 600 teachers, half are initially matched and the other half is not (they are the newcomers).

We let the parameter a for the correlation of teachers' preferences vary to capture both extremes: independent preferences and completely correlated ones. We iterate 50 times and for each iteration, we run DA*, $BE \circ$ DA* and check whether DA* matching is one-sided maximal or not and then compare the number of teachers staying at their initial position among the teachers initially matched under DAs and $BE \circ$ DA*.[32]

[32] Note that here, since all schools are acceptables, teachers rank all the schools and there are enough seats for all teachers, newcomers necessarily obtain a school in each algorithm so that there is no point of calculating those staying at their initial position (i.e. non being matched) in that case.

Table 4: Simulations results

[0,a]	DA* init	BE$\circ$DA* init
[0,0.1]	22.54	22.08
[0,0.5]	54.5	43.08
[0,15]	187.58	151.54
[0,30]	186.46	165.86
[0,60]	187.38	176.36
[0,100]	189.22	180.6
[0,1000]	190.2	189.2

For $a = 15, 30, 60$, over the 50 iterations, none of DA* matchings were two-sided maximal. For $a = 0.5$ only one was two-sided maximal and none were one-sided maximal. As for $a = 100$ only 3 were two-sided maximal and none were one-sided maximal. As correlation increases, the scope of improvement decreases. For the two extreme values, few improvements for both sides are possible, only 9 DA* matchings were not two-sided maximal for $a = 0.1$ (almost completely independant preferences), but all of them were not one-sided maximal. For almost completely correlated preferences, $a = 1000$, only 13 DA* matchings were not two-sided maximal (and 22 were not one-sided maximal).

B Proof of Theorem 3

We want to prove the following proposition.

Proposition 6 *Let φ be any selection of BE. If $\varphi \neq$ TO-BE then φ is not strategy-proof.*

Lemma 3 *Let φ be any selection of BE. Fix any profile of preferences $\succ$ and assume that $\varphi(\succ) \neq$ TO-BE$(\succ)$. Let x be the outcome of TO-BE$(\succ)$ and let y be that of $\varphi(\succ)$. There exists t s.t. $x(t) \succ_t y(t) \succ_t \mu_0(t)$.*

Proof. Let $T(x, y)$ be the set of teachers for which $x(t) \succ_t y(t) \succeq_t \mu_0(t)$. We know that x is not 1-Pareto-dominated by y (by Proposition 5) and since y is individually rational

and $x \neq y$, we must have $T(x,y) \neq \emptyset$. Proceed by contradiction and assume that for all $t \in T(x,y)$, we have $y(t) = \mu_0(t)$. Let $B := T \backslash T(x,y)$. Note that for any $t \in B$, $y(t)$ is a school initially assigned to some teacher in B. In addition, by definition, for all $t \in B$, $y(t) \succeq_t x(t)$. If there was no teacher $t \in B$ for which $y(t) \succ_t x(t)$, then we would have the following situation: y would select the initial allocation for all $t \in T(x,y)$ and would be identical to x for all $t \notin T(x,y)$. Given that $x \neq y$, we must have $x(t) \neq y(t) = \mu_0(t)$ for some $t \in T(x,y)$. Since x is individually rational, we have $x(t) \succ_t y(t) = \mu_0(t)$ for those $t \in T(x,y)$. Hence x 1-Pareto-dominates y. But all schools are also better-off under x rather than under y. Indeed, for each school s s.t. $y(s) \notin T(x,y)$, $y(s) = x(s)$ and for each school s s.t. $y(s) \in T(x,y)$, because x is individually rational on both sides, $x(s) \succeq_s y(s) = \mu_0(s)$ with a strict inequality for s satisfying $x(s) \neq y(s)$ (and such a s must exist since $x \neq y$). Thus, x is individually rational on both sides and 2-Pareto-dominates y, which is not possible given that y is an outcome of BE.

To recap, we have that for any $t \in B$, $y(t)$ is a school initially assigned to some teacher in B and for all $t \in B$, $y(t) \succeq_t x(t)$ with a strict inequality for some $t \in B$. In addition, since y is the outcome of $\varphi(\succ)$ and φ 2-Pareto-dominates the initial allocation μ_0, we must have that for all school s, $y(s) \succeq_s \mu_0(s)$. Hence, B is a two-sided blocking coalition for x, which is a contradiction since x must be a point in the two–sided Core. ■

Proof of Proposition 6. We start from a profile of preferences $\succ$ under which $\varphi(\succ) \neq$ TO-BE($\succ$) which must exist by our assumption that $\varphi \neq$ TO-BE. Given our profile of preferences $\succ$, we let the profile of preferences $\succ'$ be defined as follows. For any t, any school s other than TO-BE($\succ$)[t] are ranked as unacceptable for t under $\succ'$. We must have TO-BE($\succ$) = TO-BE($\succ'$). Now, we are in a position to prove the following lemma.

Lemma 4 *TO-BE*$(\succ') = \varphi(\succ')$.

Proof. Suppose $x :=$ TO-BE($\succ'$) $\neq \varphi(\succ') =: y$. By the above lemma, there exists t s.t. $x(t) \succ'_t y(t) \succ'_t \mu_0(t)$ which yields a contradiction, by construction of $\succ'_t$. ■

Note that TO-BE satisfies also the following property: for any profile of preferences $\succ$, for any teacher t, TO-BE($\succ$)(t) = TO-BE($\succ_{-t}, \succ'_t$)(t). This will be used in the following lemma.

Lemma 5 *If φ is strategy-proof then TO-BE$(\succ_Z, \succ'_{-Z}) = \varphi(\succ_Z, \succ'_{-Z})$ for any $Z \subseteq T$.*

Proof. Assume φ is strategy-proof. The proof is by induction on the size of Z. For $|Z| = 0$, the result is given by the previous lemma. Now, the induction hypothesis is that TO-BE$(\succ_Z, \succ'_{-Z}) = \varphi(\succ_Z, \succ'_{-Z})$ for any subset Z with $|Z| = k$. Proceed by contradiction and suppose that there is Z s.t. $|Z| = k+1$ for which $x :=$ TO-BE$(\succ_Z, \succ'_{-Z}) \neq \varphi(\succ_Z, \succ'_{-Z}) =: y$. By the first lemma above, there exists t s.t. TO-BE$(\succ_Z, \succ'_{-Z})(t) \rhd_t \varphi(\succ_Z, \succ'_{-Z})(t) \rhd_t \mu_0(t)$ where $\rhd_t = \succ'_t$ if $t \notin Z$ while $\rhd_t = \succ_t$ otherwise. If $t \notin Z$, then there is a straightforward contradiction since under $\succ'_t$ there is a single school which is ranked above $\mu_0(t)$ for teacher t. Now, assume that $t \in Z$. By the property noticed just before the statement of the lemma, we must have TO-BE$(\succ_{Z\setminus\{t\}}, \succ'_{-Z}, \succ'_t)(t) =$ TO-BE$(\succ_Z, \succ'_{-Z})(t)$ and, by our induction hypothesis, $\varphi(\succ_{Z\setminus\{t\}}, \succ'_{-Z}, \succ'_t)(t) =$ TO-BE$(\succ_{Z\setminus\{t\}}, \succ'_{-Z}, \succ'_t)(t)$. Thus, we obtain $\varphi(\succ_{Z\setminus\{t\}}, \succ'_{-Z}, \succ'_t)(t) =$ TO-BE$(\succ_{Z\setminus\{t\}}, \succ'_{-Z}, \succ'_t)(t) =$ TO-BE$(\succ_Z, \succ'_{-Z})(t) \succ_t \varphi(\succ_Z, \succ'_{-Z})(t)$ which is a contradiction with the assumption that φ is strategy-proof (indeed, at $(\succ_Z, \succ'_{-Z})$, teacher $t \in Z$ has an incentive to report $\succ'_t$ instead of $\succ_t$). ■

Taking $Z = T$ in the above lemma, given that $\varphi(\succ) \neq$ TO-BE$(\succ)$, we obtain the following corollary which completes the proof of our proposition.

Corollary 2 *φ is not strategy-proof.*

■

C Proof of Theorem 4

For a given matching μ let $\mathcal{B}_\mu$ be the set of blocking pairs of μ i.e. the pairs (t, s) s.t teacher t blocks with school s. In the next proposition, μ^k is the matching obtained at step k of the 1S-BE algorithm (whatever the selection of cycles). It confirms that the set of blocking pairs shrinks along the 1S-BE algorithm.

Proposition 7 *For all $k \geq 1$, the set of blocking pairs of μ^{k+1} is a subset of that of μ^k.*

Proof. The proof proceeds by contradiction. Assume that there is a teacher t and a school s s.t $(t,s) \in \mathcal{B}_{\mu^{k+1}}$ but $(t,s) \notin \mathcal{B}_{\mu^k}$. So $s \succ_t \mu^{k+1}(t) \succeq_t \mu^k(t)$ where the last inequality comes from the 1-Pareto domination of μ^{k+1} over μ^k by construction of 1S-BE. Since $(t,s) \notin \mathcal{B}_{\mu^k}$ and $(t,s) \in \mathcal{B}_{\mu^{k+1}}$ we should have that $\mu^k(s) \succ_s t \succ_s \mu^{k+1}(s) := t'$. So we know that t' and t were not blocking with s under μ^k and that t is preferred to t' at s so that by construction of the graph of 1S-BE at μ^k, the node $(t', \mu_k(t'))$ cannot point to the node $(\mu^k(s), s)$ contradicting the fact that t' was matched to s through a cycle between step k and $k+1$ of 1S-BE. ■

The following proposition states that the set of possible matchings of 1S-BE is included in the set of one-sided maximal matchings.

Proposition 8 *For a given instance of preferences, and a given selection of cycles, let μ be the matching obtained by 1S-BE. Then, there does not exist a matching μ' s.t μ' 1-Pareto dominates μ and $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_\mu$.*

We will prove Proposition 8 using several Lemmas. As in Lemma 1, we can exhibit "cycles of exchanges "which can be used to go from μ to μ' in the proposition. Denote these cycles of exchanges C_k, with $k = 1, \dots, K$ and where $C_k := \{(t_1^k, s_1^k), \dots, (t_{L_k}^k, s_{L_k}^k)\}$ and where $\forall l = 1, \dots, L_1 - 1$: i) $\mu(t_l^k) = s_l^k$ and ii) $\mu'(t_l^1) = s_{l+1}^1 \succ_{t_l^1} s_l^1$. So once we implement these cycles from μ, we obtain μ'. Let $\forall k = 1, \dots, K$, $\forall l = 1, \dots, L_k$, $n_l^k := (t_l^k, s_l^k)$ be a node and $C := \bigcup_{k=1}^{K} C_k$ be the set of nodes which define the different cycles C_k. As said before, note that edges in these "cycles of exchanges "between n_l^k and n_{l+1}^k in a given cycle C_k are not necessarily actual edges of the graph of 1S-BE. So we will call a 1S-BE-edge an actual edge between two nodes in the graph of 1S-BE. The following lemma exhibits a particular 1S-BE-pointing relationship between nodes in C if the matching μ' dominates the matching μ in terms of teachers' welfare and stability:

Lemma 6 *Take a matching μ' which 1-Pareto dominates a matching μ and s.t $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_\mu$. In the "cycles of exchanges "C_k previously defined, fix $(t_l^k, s_l^k) \in C_k$ for $k = 1, \dots, K$ and $l = 1, \dots, L_k$, then:*

1. *either (t_{l-1}^k, s_{l-1}^k) 1S-BE-points to (t_l^k, s_l^k).*

2. *or* $\exists k' = 1, \ldots, K$ *and* $l' = 1, \ldots, L_{k'}$ *with* $l' \neq l-1$ *if* $k' = k$ *s.t* $t_{l'}^{k'}$ *does not block with* s_l^k *under* μ *but* $s_l^k \succ_{t_{l'}^{k'}} s_{l'}^{k'}$ *and* $t_{l'}^{k'}$ *is first ranked among those who desire* s_l^k *but do not already block with* s_l^k*. It then implies that* $(t_{l'}^{k'}, s_{l'}^{k'})$ *1S-BE-points to* (t_l^k, s_l^k) *according to the second condition for an edge in 1S-BE.*

Proof. Take a node $n_l^k \in C_k$, if it does not 1S-BE-point to n_{l+1}^k then it means that $(t_l^k, s_{l+1}^k) \notin \mathcal{B}_\mu$ and that there exists a teacher t s.t $s_{l+1}^k \succ_t \mu(t)$ and: (t, s_{l+1}^k) does not block μ and t is first ranked among those who desire s_{l+1}^k and does not already block with s_{l+1}^k (cf definition of an edge above). In particular, $t \succ_{s_{l+1}^k} t_l^k$. If $(t, \mu(t)) \notin C$, then it means that we would create a blocking pair (t, s_{l+1}^k) in implementing all the cycles in C. Indeed, by construction, t does not block with s_{l+1}^k under μ. If $(t, \mu(t)) \notin C \Rightarrow \mu'(t) = \mu(t)$ so t desires s_{l+1}^k under μ' but we know that $t \succ_{s_{l+1}^k} t_l^k \in \mu'(s_{l+1}^k)$, so $(t, s_{l+1}^k) \in \mathcal{B}_{\mu'}$, a contradiction with the assumption that $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_\mu$. So necessarily, $(t, \mu(t)) \in C$. So if you fix a node n in C, either it points to its successor according to the cycles defining C, call it n' (condition 1 of the Lemma) or if it does not, as showed above, there exists another node in C which points to n' according to the second condition of the Lemma. ■

Lemma 6 states that each node in C has a particular in-edge in the graph of 1S-BE coming from another node of C. We can prove in the next lemma the existence of a particular 1S-BE-cycle between nodes of C:

Lemma 7 *Take* μ' *as defined in Lemma 6 and let* C *be the collection of nodes in the different "cycles of exchanges "*$(C_k,\, k = 1, \ldots, K)$ *which lead to* μ' *from* μ *once implemented. There exists a 1S-BE-cycle,* C^**, among nodes in* C*:* $C^* := \{(t_{l_1}^{k_1}, s_{l_1}^{k_1}), \ldots, (t_{l_M}^{k_M}, s_{l_M}^{k_M})\}$ *with* $\forall j = 1, \ldots, M$*,* $n_j^* := (t_{l_j}^{t_j}, s_{l_j}^{k_j}) \in C_{k_j}$ *and s.t:*

- *if* $n_{j+1}^* \neq n_{l_j+1}^{k_j}$*, then* $t_{l_j}^{k_j}$ *must desire* $s_{l_{j+1}}^{k_{j+1}}$ *under* μ*, not block with it and is the highest ranked among those who desire it and do not already block with it.*

Proof. Start with a node $n_1 \in C$. Then by Lemma 6, we know that there exists $n_2 \in C$ s.t: i) n_2 1S-BE-points to n_1 and ii) either n_2 is the successor of n_1 under C_k s.t $n_1 \in C_k$ or, if not, using condition 2. of Lemma 6, the teacher in n_2 is as defined in the final point of the proposition. We can then do the same operation starting from n_2 and

find a node $n_3 \in C$ which 1S-BE-points to n_2 and has the required condition. Continuing the argument, since the set of nodes C is finite, we will find a 1S-BE-cycle between nodes of C which has the required property. ■

Lemma 7 just tells that there exists a cycle under the graph of 1S-BE between nodes of C and such that, if a node does not point under this cycle to its successor as defined in the cycles of C, then it means that this node must point in the cycle of 1S-BE according to the second condition which defines the edges of 1S-BE. This existence is the one used after Theorem 4. In addition to just show the existence of a 1S-BE-cycle, the lemma exhibits a particular one with specific properties about pointing behavior of the nodes. We can now prove the Proposition 8:

Proof of Proposition 8. Assume that there exists a matching μ' which 1-Pareto dominates the matching μ obtained with 1S-BE and s.t $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\mu}$. Using Lemma 7, we know that taking the graph of 1S-BE starting from μ, there exists a cycle in this graph, contradicting that μ is the matching obtained with 1S-BE since the algorithm stops when there are no cycles left. ■

Next proposition shows that cycles in the graph of 1S-BE can be used to go from a matching μ to a matching μ' which 1-Pareto dominates it and leads to less blocking pairs in the inclusion sense.

Proposition 9 *Start from an allocation μ and take an allocation μ' s.t:*

1. *μ' 1-Pareto dominates μ*
2. *$\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\mu}$*

Then there exists a selection of cycles in the graph of 1S-BE that leads to μ' starting from μ.

Proof. Since μ' 1-Pareto dominates μ, we know that there are “cycles of exchanges ”C_k, $k = 1, \ldots, K$ which, once all implemented, lead to μ' from μ. We will use the same notations and denote by $C = \bigcup_{k=1}^{K} C_k$ the set of nodes in these cycles which lead to μ' from

μ. Since $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_\mu$ we can apply Lemma 7 to exhibit a 1S-BE-cycle, C^*, involving only nodes in C and which has the special property stated in the point of the Lemma 7. This 1S-BE-cycle, C^*, once implemented, leads to an allocation $\tilde{\mu}$. In the following, we will use the same notations as in Lemma 7 to denote nodes of C^*. The special property of Lemma 7, allows us to deduce the following remarks which can be viewed just as corollaries of Lemma 7:

Remark 1. It says that if a teacher being part of a node under C^* is not matched under $\tilde{\mu}$ to his school under μ', then it means that the edge linking his node to its successor under C^* is due to the second condition in the definition of an edge in 1S-BE. Moreover, by construction of this cycle, it means that this teacher is preferred to one teacher matched to this school under μ' (the one whose node was pointing to the node under the cycles in C). Formally, take a node $(t^{k_j}_{l_j}, s^{k_j}_{l_j}) \in C^*$, if $\tilde{\mu}(t^{k_j}_{l_j}) = s^{k_{j+1}}_{l_{j+1}} \neq s^{k_j}_{l_j+1} = \mu'(t^{k_j}_{l_j})$ then, by the construction of C^* using Lemma 7, it means that $t^{k_j}_{l_j}$ was not blocking with $s^{k_{j+1}}_{l_{j+1}}$ under μ, desired it and is the highest ranked among those who didn't block with it and desired it. In particular, $t^{k_j}_{l_j} \succ_{s^{k_{j+1}}_{l_{j+1}}} t^{k_{j+1}}_{l_{j+1}-1} \in \mu'(s^{k_{j+1}}_{l_{j+1}})$.

Remark 2. Note that in C^*, if a node points to its successor according to the first condition in the definition of an edge (i.e. the teacher is blocking), then it means that this successor is the same one as the one defined in the cycles which form C and are used to go from μ to μ'. Formally, if the 1S-BE-edge, linking two consecutive nodes in C^*, $n^*_j = (t^{k_j}_{l_j}, s^{k_j}_{l_j})$ and $n^*_{j+1} = (t^{k_{j+1}}_{l_{j+1}}, s^{k_{j+1}}_{l_{j+1}})$, is due to the first condition i.e. forming a blocking pair in the definition of the edges of the graph of 1S-BE at μ. Then n^*_{j+1} is the successor of n^*_j in the "cycle of exchanges "C_{k_j}. So it means that $(t^{k_{j+1}}_{l_{j+1}}, s^{k_{j+1}}_{l_{j+1}}) = (t^{k_j}_{l_j+1}, s^{k_j}_{l_j+1})$ and so that $\tilde{\mu}(t^{k_j}_{l_j}) = \mu'(t^{k_j}_{l_j}) = s^{k_j}_{l_{j+1}}$.

Note first, that μ' weakly 1-Pareto dominates $\tilde{\mu}$. Indeed, by construction of the 1S-BE-cycle, take a teacher t:

- either $\mu(t) = \mu'(t)$ and in that case $\tilde{\mu}(t) = \mu'(t)$ since $(t, \mu(t)) \notin C$ so cannot be part of the 1S-BE-cycle. So t is indifferent between μ' and $\tilde{\mu}$.

- either $\tilde{\mu}(t) = \mu'(t)$. So in that case, t is also indifferent between μ' and $\tilde{\mu}$.
- either $\mu(t) \neq \tilde{\mu}(t) \neq \mu'(t)$. In that case, applying Remark 1, under μ, t did not block with $\tilde{\mu}(t)$, i.e. $(t, \tilde{\mu}(t)) \notin \mathcal{B}_\mu$, and was the highest ranked teacher among those who desired $\tilde{\mu}(t) := s$ but did not block with it. In particular, $t \succ_s t'$ for some $t' \in \mu'(s) \backslash \mu(s)$ (the one whose node was pointing to a node with s under the cycles of C). If $s \succ_t \mu'(t)$, it would mean that $(t, s) \in \mathcal{B}_{\mu'}$ (since t has a justified envy at s against t') but we know that $(t, s) \notin \mathcal{B}_\mu$ and $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_\mu$, a contradiction.

Then, if we can prove that $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\tilde{\mu}}$, the proof would be done since we can apply the same reasoning as above, find a 1S-BE-cycle, but this time, starting from $\tilde{\mu}$ instead of μ.

So take $(t, s) \in \mathcal{B}_{\mu'}$. Note that by 1-Pareto domination, we have that $s \succ_t \mu'(t) \succeq_t \tilde{\mu}(t) \succeq_t \mu(t)$. Then, we only need to check that t is preferred to the teacher matched to $\tilde{\mu}$, there are several cases:

1. if $\mu'(s) = \mu(s)$: necessarily $\tilde{\mu}(s) = \mu'(s) = \mu(s)$ since by construction, teachers who do not move between μ and μ' cannot be in C^* and so do not move from μ to $\tilde{\mu}$ either. So $(t, s) \in \mathcal{B}_{\tilde{\mu}}$ since the teacher against whom t has a justified envy under μ' is still matched to s under $\tilde{s}$.
2. if $\mu'(s) \neq \mu(s)$:
   (a) if $\tilde{\mu}(s) = \mu'(s) \neq \mu(s)$: same argument than before, so $(t, s) \in \mathcal{B}_{\tilde{\mu}}$.
   (b) $\tilde{\mu}(s) = \mu(s) \neq \mu'(s)$: if $(t, s) \notin \mathcal{B}_{\tilde{\mu}}$ then $(t, s) \notin \mathcal{B}_\mu$, a contradiction with $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_\mu$.
   (c) $\tilde{\mu}(s) \neq \mu(s) \neq \mu'(s)$: Since $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_\mu$, we know that $(t, s) \in \mathcal{B}_\mu$. Assume $(t, s) \notin \mathcal{B}_{\tilde{\mu}}$. So it means that the teacher against whom t had a justified envy under μ left s under $\tilde{\mu}$, so was part of some nodes in the cycle C^*. So take $t' = \mu(s)$ s.t $t \succ_s t'$ (we know that it exists at least one since $(t, s) \in \mathcal{B}_\mu$), we know that the node $(t', \mu(t'))$ must be part of C^*. So under C^*, there existed a node (t'', s'') which 1S-BE-pointed to (t', s), note that $t'' \neq t$. Indeed, we know that $(t, \mu(t))$ was 1S-BE-pointing to (t', s) due to the fact that he had a justified

envy against t' at s (first condition of the definition of an edge). We know by Remark 2, that all the nodes which 1S-BE-point under C^* due to the blocking pair condition must point to their successor according to the cycles in C. So if $t'' = t$ it would mean that $s = \mu'(t)$ and so (t, s) would not be a blocking pair under μ', a contradiction to our initial assumption. Since $(t, s) \notin \mathcal{B}_{\tilde{\mu}}$ by assumption, necessarily $t'' \succ_s t \succ_s t'$ (otherwise t would block with s under $\tilde{\mu}$). So t'' had a justified envy at s against t. But by Remark 2, it necessarily means that $s = \mu'(t'')$. So we have that $(t, s) \notin \mathcal{B}_{\mu'}$, a contradiction with our initial assumption.

So we showed that $\mathcal{B}_{\tilde{\mu}} \subseteq \mathcal{B}_{\mu'}$, which completes the proof. ■

D Proof of Theorem 5

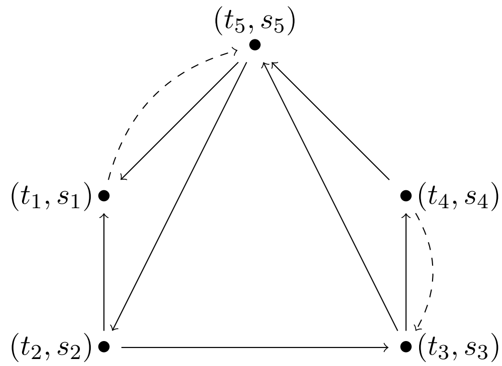
In order to prove this result, we exhibit an instance where, irrespective of which cycle one selects in the graphs associated with 1S-BE , one teacher will gain by misreporting his preferences. Assume that there are five teachers $t_1, \ldots, t_5$ and five schools $s_1, \ldots, s_5$. Teachers' and schools' preferences are given as follows:

$$
\begin{array}{llllllllll}
\succ_{t_1}: & s_5 & s_1 & & \qquad \succ_{s_1}: & t_5 & t_2 & t_1 & & \\
\succ_{t_2}: & s_1 & s_3 & s_2 & \qquad \succ_{s_2}: & t_5 & t_2 & & & \\
\succ_{t_3}: & s_4 & s_5 & s_3 & \qquad \succ_{s_3}: & t_3 & t_2 & t_4 & & \\
\succ_{t_4}: & s_5 & s_3 & s_4 & \qquad \succ_{s_4}: & t_3 & t_4 & & & \\
\succ_{t_5}: & s_2 & s_1 & s_5 & \qquad \succ_{s_5}: & t_4 & t_2 & t_5 & t_3 & t_1
\end{array}
$$

We let $\succ := (\succ_{t_1}, \ldots, \succ_{t_5})$. The initial assignment is given by:

$$
\mu = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 & t_5 \\ s_1 & s_2 & s_3 & s_4 & s_5 \end{pmatrix}
$$

Starting from the initial assignment, the solid arrows in the graph below correspond to the graph associated with 1S-BE.



We added dashed arrows from one vertex to the other if the teacher in the origin of the arrow prefers the school in the pointed vertex. Theses arrows are not actual arrows of the graph associated with 1S-BE and so cannot be used to select a cycle, there are just here to facilitate the understanding of the argument.

When $\succ$ is submitted, there are two possible choices of cycles in the graph:

- A "large"cycle given by: $(t_2, s_2) \rightarrow (t_3, s_3) \rightarrow (t_4, s_4) \rightarrow (t_5, s_5) \rightarrow (t_2, s_2)$. Denote this cycle by $\bar{C}$.
- A "small"cycle given by: $(t_2, s_2) \rightarrow (t_3, s_3) \rightarrow (t_5, s_5) \rightarrow (t_2, s_2)$. Denote this cycle by $\underline{C}$.

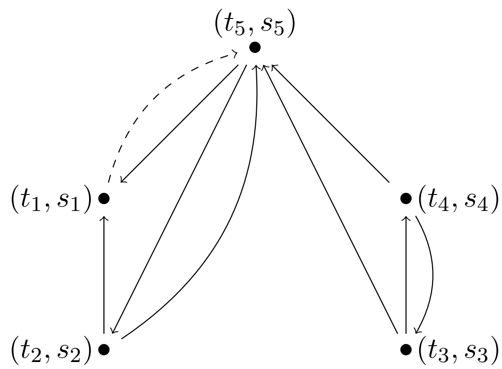
So we decompose the analysis for these two cases.

Case A: Under $\succ$, $\bar{C}$ is chosen:

Once this cycle is cleared, there are no cycles left in the graph associated with 1S-BE and the final matching of 1S-BE is given by:

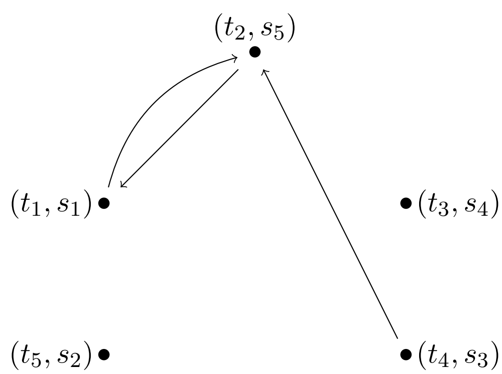
$$\bar{\mu} = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 & t_5 \\ s_1 & s_3 & s_4 & s_5 & s_2 \end{pmatrix}$$

Now, assume that teacher t_2 reports the following preference relation: $\succ'_{t_2}$: s_1, s_5, s_2 while others report according to $\succ$. Under this profile, starting from the initial assignment, the graph associated with 1S-BE is:



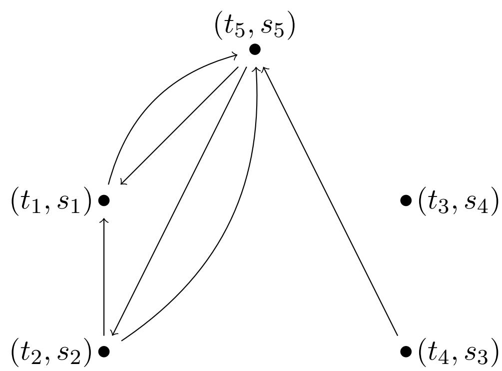
Now, there are two possible choices of cycles.

Case A.1: The cycle chosen is $(t_2, s_2) \leftrightarrows (t_5, s_5)$. Once carried out, the graph associated with 1S-BE starting from the new matching is:



Clearly, the only possible choice is to match teacher t_2 to school s_1. Hence, t_2 obtains his most preferred school under $\succ_{t_2}$ and so we exhibited a profitable misreport.

Case A.2: The cycle chosen is $(t_4, s_4) \leftrightarrows (t_3, s_3)$. Once carried out, the graph associated with 1S-BE starting from the new matching is:



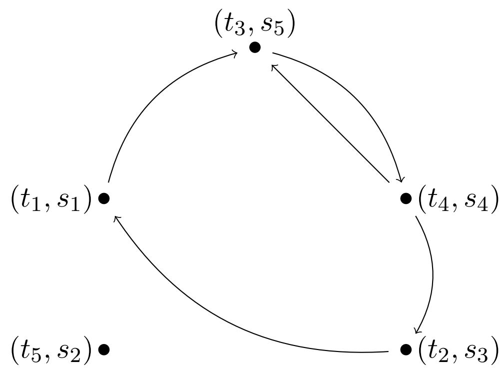
In this graph, there are three possible choices of cycles:

1. $(t_2, s_2) \to (t_1, s_1) \to (t_5, s_5) \to (t_2, s_2)$: in that case t_2 is matched to s_1 and so, again, we identified a profitable misreport.
2. $(t_2, s_2) \leftrightarrows (t_5, s_5)$: Once cleared, the only cycle that is left is $(t_1, s_1) \leftrightarrows (t_2, s_5)$ and so t_2 will be matched to s_1 leading to a successful manipulation.
3. $(t_1, s_1) \leftrightarrows (t_5, s_5)$: Once cleared, since t_5 prefers s_2 to s_1 there is a unique cycle left which is $(t_5, s_1) \leftrightarrows (t_2, s_2)$. Once again the manipulation of t_2 is successful.

Thus, we have shown that when choosing cycle $\overline{C}$ under the the profile $\succ$, teacher t_2 has a profitable misreport irrespective of the possible selections of cycles performed after t_2's deviation. Let us now move to the other case.

Case B: Under $\succ$, $\underline{C}$ is chosen:

Once this cycle is carried out, the graph associated with 1S-BE starting from the new matching is:



There are two possible choices of cycles.

Case B.1: $(t_3, s_5) \leftrightarrows (t_4, s_4)$ is chosen. Then the matching obtained is the same as the one obtained with the choice of cycle $\bar{C}$. So we come back to *Case A* and we know that t_2 has a successful misreport.

Case B.2: $(t_1, s_1) \to (t_3, s_5) \to (t_4, s_4) \to (t_2, s_3) \to (t_1, s_1)$ is chosen. In that case, assume that t_4 submits the following preferences: $\succ'_{t_4}$: s_5, s_4. The graph associated with 1S-BE starting from the initial assignment is the same as the one under truthful reports (note that, although these are not the arrow of the graph of 1S-BE, the dashed arrow from (t_4, s_4) disappears). So, again, we are left with a choice between cycle $\bar{C}$ and $\underline{C}$.

1. If carry out $\underline{C}$, the graph starting from the new matching will be given by the graph just above except that now (t_4, s_4) does not point to (t_2, s_3) anymore. Hence, we can only pick cycle $(t_3, s_5) \leftrightarrows (t_4, s_4)$ and so t_4 obtains his best school and we identified a profitable misreport for teacher t_4.

2. If we choose $\bar{C}$, we already know that we end up with matching $\bar{\mu}$ as defined above. So, here again, t_4 obtains his best school s_5 and the manipulation is also a success.

To sum up, we have shown that for each possible selection of cycles under 1S-BE, there is a teacher who has a profitable misreport. Thus, no selection of the 1S-BE algorithm is strategy-proof, as was to be shown.

E Proof of Theorem 6, 7 and 8

E.1 Preliminaries in random graph

In the sequel, we will exploit two standard results in random graph theory that are stated in this section. It is thus worth introducing the relevant model of random graph. A **graph** $G(n)$ consists in n vertices, V, and edges $E \subseteq V \times V$ across V. A random graph $G(n,p)$, where $p \in (0,1)$, is a graph with n vertices V in which each pair $(v_1, v_2) \in V \times V$ is linked by an edge with probability p independently (of edges created for all other pairs).

A **perfect matching** of $G(n)$ is a subset E' of E such that each node in V is contained in a single edge of E'. A perfect matching may not always exist. However, the following standard result states that in large random graphs it is very likely that such a matching exists.

Lemma 8 (Erdös-Rényi) *Fix a random graph $G(n,p)$. The probability that there is a perfect matching in a realization of $G(n,p)$ tends to 1 as $n \to \infty$.*

The second important technical result will about so called independent sets. An **independent set** is $\bar{V} \subseteq V$ such that any $(i,j) \in \bar{V} \times \bar{V}$, (i,j) is not in E.

Lemma 9 *Let Z be an independent set in a random graph $G(n,p)$. Then,*

$$\Pr\left\{|Z| \geq \frac{2\log n}{\log\frac{1}{1-p}} + 1\right\} \to 0 \text{ as } n \to \infty.$$

Proof. Let Z_a be the number of independent sets of size a in G. For $a = \frac{2\log n}{\log\frac{1}{1-p}} + 1$, we must have

$$\Pr\left\{|Z| \geq \frac{2\log n}{\log\frac{1}{1-p}} + 1\right\} \leq \Pr\{Z_a \geq 1\} \leq \mathbb{E}[Z_a] = \binom{n}{a}(1-p)^{\binom{a}{2}}.$$

The computation of $\mathbb{E}[Z_a]$ follows from the following argument. A subset of nodes $A \subset V$ forms an independent set, if there is no edge between every pair of nodes $u, v \in A$. Suppose A has size a. Since the probability of an edge is p, the probability that a given node set

A forms an independent set is $(1-p)^{\binom{a}{2}}$. There are $\binom{n}{a}$ different ways of choosing a node subset $A \subset V$ of size a. Hence, the expected number of independent sets of size a is

$$\binom{n}{a}(1-p)^{\binom{a}{2}}.$$

Note that since $a = \frac{2\log n}{\log\frac{1}{1-p}} + 1 = -2\log_{1-p} n + 1$, we have

$$(1-p)^{\binom{a}{2}} = (1-p)^{\frac{a(a-1)}{2}} = ((1-p)^{\frac{(a-1)}{2}})^a \leq ((1-p)^{-\log_{1-p} n})^a = n^{-a}.$$

In addition to this, we have

$$\binom{n}{a} \leq \frac{n^a}{a!}.$$

Hence, for $a = \frac{2\log n}{\log\frac{1}{1-p}} + 1$, we obtain that

$$\Pr\{Z_a \geq 1\} \to 0 \text{ as } n \to \infty.$$

■

Remark 1 *This result is related to Matula (1970, 1972, 1976), Grimmett and McDiarmid (1975), Bollobàs and Erdös (1976), Chvàtal (1977).*

E.2 Proof of Theorem 7

In the sequel, we let T_k be $\mu_0(S_k)$ where μ_0 is the initial allocation. We will prove the following result which implies the first part of Theorem 7.

Proposition 10 *Consider any selection φ of the BE-algorithm. Fix any k. Let $\bar{T}_k := \{t \in T_k \,|\varphi(t) \neq \mu_0(t)\}$. We have*

$$\frac{|\bar{T}_k|}{|T_k|} \xrightarrow{p} 1.$$

Proof of Proposition 10. Fix an arbitrary k and fix $\epsilon > 0$. We define a random graph with $\{(t, \mu_0(t))_{t\in T_k}$ as the set of vertices. An edge between $(t, \mu_0(t))$ and

$(t', \mu_0(t'))$ is added if and only if $\xi_{t\mu_0(t')} > 1 - \epsilon$ and $\xi_{t'\mu_0(t)} > 1 - \epsilon$ and $\eta_{t'\mu_0(t)} > 1 - \epsilon$ and $\eta_{t\mu_0(t')} > 1 - \epsilon$. Then, in the random graph, each edge between $(t, \mu_0(t))$ and $(t', \mu_0(t'))$ is added independently with probability $\epsilon^4 \in (0, 1)$. Then, let $\hat{T}_k := \{t \in T_k \,|\varphi(t) = \mu_0(t)$ and $U_t(\mu_0(t)) \leq u_k + 1 - \varepsilon$ and $V_{\mu_0(t)}(t) \leq 1 - \varepsilon\}$. It must be that $\{(t, \mu_0(t))_{t \in \hat{T}_k}$ is an independent set, or else if there is an edge $(t, \mu_0(t)), (t', \mu_0(t'))$ for some realization of the random graph, then

$$U_t(\mu_0(t')) > u_k + 1 - \epsilon \geq U_t(\mu_0(t)) = U_t(\varphi(t)) \text{ and } V_{\mu_0(t')}(t) > 1 - \epsilon \geq V_{\mu_0(t')}(t') = V_{\mu_0(t')}(\varphi(\mu_0(t')))$$

and similarly,

$$U_{t'}(\mu_0(t)) > u_k + 1 - \epsilon \geq U_{t'}(\mu_0(t')) = U_{t'}(\varphi(t')) \text{ and } V_{\mu_0(t)}(t') > 1 - \epsilon \geq V_{\mu_0(t)}(t) = V_{\mu_0(t)}(\varphi(\mu_0(t))).$$

Put in another way, both $(t, \mu_0(t'))$ and $(t', \mu_0(t))$ block φ. Since, by definition, under φ, t is assigned $\mu_0(t)$ and t' is assigned $\mu_0(t')$, this means that there are still cycles in the graph associated with BE when starting from the assignment given by φ which contradicts the fact that φ is a selection of BE.

Now, we can use Lemma 9 to get that $\Pr\left\{|\hat{T}_k| \geq \frac{2\log(|T_k|)}{\log \frac{1}{1-p}} + 1\right\} \to 0$ as $n \to \infty$ and thus $\frac{|\hat{T}_k|}{|T_k|} \xrightarrow{p} 0$ as $n \to \infty$. Setting $\tilde{T}_k := \{t \in T_k \,|U_t(\mu_0(t)) \leq u_k + 1 - \varepsilon$ and $V_{\mu_0(t)}(t) \leq 1 - \varepsilon\}$, we have

$$\frac{|\hat{T}_k|}{|T_k|} = \frac{|\bar{T}_k^c \cap \tilde{T}_k|}{|T_k|} = \frac{|\bar{T}_k^c \backslash \tilde{T}_k^c|}{|T_k|} \geq \frac{|\bar{T}_k^c|}{|T_k|} - \frac{|\tilde{T}_k^c|}{|T_k|}.$$

We know that for the left hand-side above : $\frac{|\hat{T}_k|}{|T_k|} \xrightarrow{p} 0$ as $n \to \infty$. By the law of large numbers, $\frac{|\tilde{T}_k^c|}{|T_k|} \xrightarrow{p} 1 - (1 - \varepsilon)^2$ which can be made arbitrarily close to 0 given that $\epsilon > 0$ is arbitrary. Hence, we obtain that $\frac{|\bar{T}_k^c|}{|T_k|} \xrightarrow{p} 0$ as $n \to \infty$, as was to be proved.

Let us now move to the other part of Theorem 7. We have to show that there exists a selection of BE which is asymptotically teacher-efficient, asymptotically school-efficient and asymptotically stable. Note that in our environment asymptotic school-efficiency implies asymptotic stability. Hence, the following proposition is enough for this purpose.

Proposition 11 *Fix any k and any $\varepsilon > 0$. Let $\bar{T}_k := \{t \in T_k \,|U_t(\text{TO-BE}(t)) \geq u_k + 1 - \varepsilon\}$ and $\bar{S}_k := \{s \in S_k \,|V_s(\text{TO-BE}(s)) \geq 1 - \varepsilon\}$. We have*

$$\frac{|\bar{T}_k|}{|T_k|} \xrightarrow{p} 1 \text{ and } \frac{|\bar{S}_k|}{|S_k|} \xrightarrow{p} 1$$

Proof of Proposition 11. To be completed.

E.3 Proof of Theorem 8

Recall that T_k stands for $\mu_0(S_k)$ where μ_0 is the initial allocation. We will prove the following result.

Proposition 12 *Fix any k and any $\varepsilon > 0$. Let $\bar{T}_k := \{t \in T_k \,|U_t(\text{TO-BE}(t)) \geq u_k + 1 - \varepsilon\}$. We have*

$$\frac{|\bar{T}_k|}{|T_k|} \xrightarrow{p} 1.$$

Proof of Proposition 12. Recall that TO-BE is in the two-sided core. In particular, this implies that there is no pair of teachers t and t' so that $\mu_0(t') \succeq_t$ TO-BE(t), $\mu_0(t) \succeq_{t'}$ TO-BE(t') (with a strict preference for either t or t'), $t' \succeq_{\mu_0(t)} t$ and $t \succeq_{\mu_0(t')} t'$. Fix an arbitrary k and let E be the event that the fraction of schools $s \in S_k$ s.t. $\eta_{\mu_0(s)s} \leq 1 - \delta$ is greater than $1 - 2\delta$ where $\delta \in (0, 1)$. By the law of large numbers, we have

$$\frac{1}{|S_k|} \sum_{s \in S_k} \mathbf{1}_{\{\eta_{\mu_0(s)s} \leq 1 - \delta\}} \xrightarrow{p} 1 - \delta.$$

Thus, $\Pr(E) \to 1$. Let $T_k^0 := \{t \in T_k \,\Big|\, \eta_{t\mu_0(t)} \leq 1 - \delta\}$.

In the sequel, we condition on event E and we fix a realization of $\left\{\eta_{\mu_0(s)s}\right\}_{s \in S}$ compatible with E. Observe that T_k^0 is non-random once this has been fixed and note that conditional on these, individuals' preferences are still drawn according to the same distribution (as in the unconditional case) and for $t \neq \mu_0(s)$, η_{ts} is also still drawn according to the same distribution. We further observe that, because that event E holds, $\frac{|T_k^0|}{|T_k|} \geq 1 - 2\delta$ and hence $\left|T_k^0\right|$ goes to infinity as $n \to \infty$. We define a random graph with $\{(t, \mu_0(t))\}_{t \in T_k^0}$ as the set of vertices. An edge between $(t, \mu_0(t))$ and $(t', \mu_0(t'))$ is added if and only if $\xi_{t\mu_0(t')} > 1 - \epsilon$ and $\xi_{t'\mu_0(t)} > 1 - \epsilon$ and $\eta_{t'\mu_0(t)} \geq \eta_{t\mu_0(t)}$ and $\eta_{t\mu_0(t')} \geq \eta_{t'\mu_0(t')}$. Then, in the

random graph, each edge between $(t, \mu_0(t))$ and $(t', \mu_0(t'))$ is added independently with probability at least $\epsilon^2\delta^2 \in (0,1)$. Now, let $\bar{T}_k^0 := \{t \in T_k^0 \,|U_t(\text{TO-BE}(t)) \leq u_k + 1 - \varepsilon\}$. It must be that $\bar{T}_k^0$ is an independent set, or else if there is an edge $(t, t') \in \bar{T}_k^0 \times \bar{T}_k^0$ for some realization of the random graph, then

$$U_t(\mu_0(t')) > u_k + 1 - \epsilon \geq U_t(\text{TO-BE}(t)) \text{ and } U_{t'}(\mu_0(t)) > u_k + 1 - \epsilon \geq U_{t'}(\text{TO-BE}(t')).$$

In addition, $V_{\mu_0(t)}(t') = \eta_{t'\mu_0(t)} \geq \eta_{t\mu_0(t)} = V_{\mu_0(t)}(t)$ and $V_{\mu_0(t')}(t) = \eta_{t\mu_0(t')} \geq \eta_{t'\mu_0(t')} = V_{\mu_0(t')}(t')$ and so TO-BE is blocked by a coalition of size two, a contradiction. Now, we can use Lemma 9 to get that $\Pr\left\{|\bar{T}_k^0| \geq \frac{2\log(|T_k|)}{\log\frac{1}{1-p}} + 1\right\} \to 0$ as $n \to \infty$ and thus $\frac{|\bar{T}_k^0|}{|T_k^0|} \xrightarrow{p} 0$ as $n \to \infty$. Now, since $\bar{T}_k^c = \bar{T}_k^0 \cup \{t \in T_k \backslash T_k^0 \,|U_t(\text{TO-BE}(t)) \leq u_k + 1 - \varepsilon\}$ we must have

$$\frac{|\bar{T}_k^c|}{|T_k|} \leq \frac{|\bar{T}_k^0| + |T_k \backslash T_k^0|}{|T_k|} \leq \frac{|\bar{T}_k^0|}{|T_k^0|} + 2\delta$$

Hence, given that $\frac{|\bar{T}_k^0|}{|T_k^0|} \xrightarrow{p} 0$, we must have that with probability going to 1 as n goes to infinity, $\frac{|\bar{T}_k^c|}{|T_k|} \leq 3\delta$ and so $\frac{|\bar{T}_k|}{|T_k|} \geq 1 - 3\delta$.

To recap, given event E and any realization of $\left\{\eta_{\mu_0(s)s}\right\}_{s \in S}$, we have $\frac{|\bar{T}_k|}{|T_k|} \geq 1 - 3\delta$ with probability going to 1 as $n \to \infty$. Since the realization of $\left\{\eta_{\mu_0(s)s}\right\}_{s \in S}$ is arbitrary, we obtain that given event E, $\frac{|\bar{T}_k|}{|T_k|} \geq 1 - 3\delta$ with probability going to 1 as $n \to \infty$. Since $\Pr(E) \to 1$ as $n \to \infty$, we get that $\frac{|\bar{T}_k|}{|T_k|} \geq 1 - 3\delta$ with probability going to 1 as $n \to \infty$. Since $\delta > 0$ is arbitrary small, we obtain $\frac{|\bar{T}_k|}{|T_k|} \xrightarrow{p} 1$ as $n \to \infty$, as claimed.

Remark 2 *The statement is related to that of (Che and Tercieux, 2015b, Theorem 1). However, since TO-BE is not Pareto-efficient, their proof/argument does not apply.*

Remark 3 *The argument relies on the fact that TO-BE is not blocked by any coalition of size* 2*. Hence, the result applies beyond the TO-BE mechanism and applies to any mechanism which cannot be blocked by any coaltion of size* 2*.*

F Many-to-one Extensions

We provide below the extensions of BE and 1S-BE to the many-to-one framework. So now, each school may have multiple seats. As before, we assume that all the teachers are

initially matched to a school and that all seats are initially occupied by a teacher. As before let μ_0 be the initial matching.

The Block Exchange Algorithm

The main difference is that now, blocking with a school does not necessarily means that a teacher is preferred to a given matched one in this school. To keep the idea of not hurting any school, we so have to allow a node to point to another one only if the teacher of the former is preferred to the teacher of the latter by the corresponding school.

- **Step** 0 : set $\mu(0) := \mu_0$.

- **Step** $k \geq 1$: Given $\mu(k-1)$, let the teachers and their assignments stand for the vertices of a directed graph where for each pair of nodes (t, s) and (t', s'), there is an edge $(t, s) \longrightarrow (t', s')$ if and only if teacher t has a justified envy against teacher t' at s' i.e. he prefers s' to its match s and is preferred by s' to t'. If there is no cycle, then return $\mu(k-1)$ as the outcome of the algorithm. Otherwise, select a cycle in this directed graph. For each edge $(t, s) \longrightarrow (t', s')$ in the cycle, assign teacher t to school s'. Let $\mu(k)$ be the matching so obtained. Go to step $k+1$.

The 1-Sided Block Exchange Algorithm

In order to keep the property that the graph of 1S-BE is a supergraph of the one of BE, we need to build on the previous extension of BE to define the extension of 1S-BE.

- **Step** 0 : set $\mu(0) := \mu_0$.

- **Step** $k \geq 1$: Given $\mu(k-1)$, let the teachers and their assignments stand for the vertices of a directed graph where for each pair of nodes (t, s) and (t', s'), there is an edge $(t, s) \longrightarrow (t', s')$ if and only if either (1) teacher t has a justified envy against t' at s'; or (2) t desires s' and t is ranked first by s' among teachers who both desire s' and do not block with c'[33]. If there is no cycle, then return $\mu(k-1)$ as the outcome

[33]Note that here, teacher t may block with s' since, in many-to-one, not having a justified envy against a particular teacher at a school does not imply not to block with the school. So he has to be compared with teachers desiring the school but not blocking with it.

of the algorithm. Otherwise, select a cycle in this directed graph. For each edge $(t, s) \longrightarrow (t', s')$ in the cycle, assign teacher t to school s'. Let $\mu(k)$ be the matching so obtained. Go to step $k + 1$.

G Empirical results

Method

For the empirical part of the analysis, we decided not to focus on the second phase of the assignment because reported preferences seem to be less reliable. First, teachers are restricted to rank up to 20 schools and it is well known in the literature (Haeringer and Klijn (2009)) that such constraint gives rise to strategic reports. Second, rather than ranking only schools, every element in teachers' ranking can be a school or a larger geographic area such as a city, a group of cities, a department or the entire region. As a result, we prefer using the data on the regional assignment.

During the first phase of the assignment, 25 067 teachers participate. We keep only the 49 subjects containing more than 10 teachers asking for a transfer : this restrict the sample to 20 808 teachers. We also remove from the sample all couples (1579 teachers), for which an additional constraint should be considered in the assignment. Finally, only teachers which have an initial assignment are kept in the sample : the empirical analysis is done on 10579 teachers.

Deterioration of regions' welfare

In order to evaluate welfare losses from regions, we define a new concept: a region is said to be **weakly worse off** under a matching μ' than under a matching μ if it is not strongly better off. Formally, a region is weakly worse off if there is a rank k s.t the k-th best matched teacher under μ' has a lower priority than the k-th best matched teacher under μ. Note that this definition does not exclude the possibility of improving the other seats. Table 5 shows that on average, 4.08% of the regions are weakly worse off under DA* than under their initial allocation, while BE and TO-BE do not hurt any region.

Being weakly worse off provides information on the seats whose priority has been

Table 5: Percentage of weakly worse off regions

	DA^*	TO-BE	BE(Init)	$BE(DA^*)$	1S-BE(Init)	$1S\text{-}BE(DA^*)$
Compared to:						
Initial assignment	4.08	0.00	0.00	4.08	15.87	13.43
DA^*	0.00	11.65	11.26	0.00	19.55	12.90

† Note: For each mechanism, we use the iteration leading to the highest movement. There are 31 regions in the sample. For each of the 49 subjects, we computed the number of weakly worse off regions and divided it by 49×31 to obtain the average percentage. The upper line of the table reports percentages of weakly worse off regions compared to their initial assignment. The bottom line reports percentages of weakly worse off regions compared to their assignment under DA^*.

deteriorated, but it does not give any information on the seats which are assigned a teacher with a higher priority. A region may be considered as weakly worse off because one of its seats is assigned a teacher with a lower priority, but all other seats might be assigned teachers with a higher priority. To take this into account, we compute in each subject and region, the number of seats assigned a better teacher minus the number of seats assigned a teacher with a lower priority.[34] Only regions whose match changes from their initial one are considered. Table 6 reports the distribution of the results (expressed in average percentage of regions). The upper line (-100%) reports the average percentage of regions having all positions deteriorated in net. In opposition, to bottom line (100%) corresponds to the percentage of regions having all positions improved in net. An interesting result is that the average percentage of regions whose positions are improved (in net) is higher than the percentage of regions whose positions are deteriorated (in net again): 23.07 % vs. 9.15 %.

Empirical results on TTC

As reported in table 1, TTC leads to a better cumulative distribution in terms of ranks obtained by teachers (in a sense of first order stochastic dominance). However, it leads to

[34] As for the strongly better and weakly worse off concepts we compare 'seats' in comparing the k-th best ranked teachers of each matching.

Table 6: Distribution of the percentage of regions having a net welfare improvement (compared to their initial assignment).

Net percentage of positions	DA*	TO-BE	BE(Init)	BE(DA*)	1S-BE(Init)	1S-BE(DA*)
-100/-91%	0.26	0.00	0.00	0.26	1.71	1.12
-90/-71%	0.07	0.00	0.00	0.07	0.59	0.33
-70/-51%	0.20	0.00	0.00	0.20	0.99	1.05
-50/-31%	0.66	0.00	0.00	0.59	2.24	2.11
-30/-1%	1.58	0.00	0.00	1.05	3.62	2.83
0%	0.20	0.00	0.00	0.20	1.51	1.25
1/29%	3.03	5.07	4.28	4.61	3.88	3.82
30/49%	2.50	4.21	4.67	3.62	3.62	3.09
50/69%	3.29	5.53	5.99	5.66	4.94	5.13
70/89%	2.76	3.29	4.41	4.08	2.63	3.16
90/100%	2.24	5.33	6.25	7.44	7.97	9.15
Total	16.79	23.44	25.61	27.78	33.71	33.05

† Note: Sample: Regions whose match changes from their initial assignment. For each mechanism, we use the iteration leading to the highest movement.
For each of the 49 subjects*31 regions, we computed the number of positions being assigned a teacher with a higher priority, to which we substracted the number of positions being assigned a teacher with a lower priority. Then, for each subjects*regions, the net total was divided by the total number of positions to obtain the percentage of positions being improved in net terms. Finally, the total number of regions considered has been divided by 49×31 to obtain the average percentage of regions. For instance, on average, under DA* 3.29 % of the regions have between 1 and 29 % of their seats assigned a teacher with a higher priority (in net).

less movement than 1S-BE: 120 more teachers stay at their initial allocation. Concerning the number of blocking pairs, surprisingly less teachers are part of a blocking pair under TTC than under BE and TO-BE (4292 teachers vs. 3948 for BE). However, more teachers are part of a BP under TTC than under 1S-BE (Table 3). These differences make sense since TTC is designed to focus only on teachers' welfare contrary to 1S-BE which, in addition, takes into account fairness. One of the main drawbacks of TTC is that it may create new blocking pairs compared to the initial allocation. This would not be possible with any of our mechanisms. Indeed there are 237 teachers who are not part of a blocking

pair under the initial allocation but are under TTC. Comparing TTC with TO-BE for instance, highlights the trade-off existing between the potential creation of new blocking pairs and the potential decrease of the *overall number* of blocks. Some individuals have to be sacrificed in order to decrease the total envy. This means that some teachers will feel justified envy under TTC while they were not under the initial assignment or TO-BE.

Finally, the last important drawback of TTC (especially when compared to BE and TO-BE) is that it may seriously hurt regions' welfare side. Figure 1 plots the statistics presented in Table 6 : the distribution of the percentage of regions having a net welfare improvement. TTC hurts (in net terms) significantly more regions. Especially more regions can see almost all their seats being worse off.

Figure 1: Distribution of the percentage of regions having a net welfare improvement.