The Design of Teacher Assignment: Theory and Evidence*

Julien Combe† Olivier Tercieux‡ Camille Terrier§

June 10, 2016

Abstract

The (re-)assignment of teachers to schools is a central issue in education policies. In several countries, this assignment is managed by a central administration which faces a key constraint: making 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. We show that this mechanism fails to be efficient in a strong sense: one can reassign teachers in a way which makes both teachers and schools "better-off". In addition, this reassignment increases "fairness"by shrinking the set of blocking pairs. To go around this weakness, we characterize the class of mechanisms which cannot be improved upon in terms of both efficiency and fairness. Our main theoretical finding shows that this class contains an essentially unique strategy-proof mechanism. We refine these results in two ways. First, we consider a random environment where preferences and schools' rankings are drawn randomly from a rich class of distributions and show that when the market becomes large, our mechanism "perform quantitatively better" than the modified DA in terms of utilitarian efficiency and number of blocking pairs. Second, we use a rich dataset on teachers' applications to transfer in France to empirically assess the extent of potential gains associated with the adoption of our mechanisms. These empirical results confirm both the poor performance

*We are grateful to Francis Bloch, Yeon-Koo Che, Jinwoo Kim, Jacob Leshno, Parag Pathak, Juan Pereyra and seminar participants at Columbia University, LSE, PSE, Seoul National University for helpful comments. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the ≪ Investissements d'avenir ≫ program (reference : ANR-10-EQPX-17 – Centre d'accès sécurisé aux données – CASD).

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

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

§Paris School of Economics, CEP (LSE) and IZA. Email: c.terrier@lse.ac.uk.

of the variation of the DA mechanism, and the significant improvement our alternative mechanisms bring in terms of both efficiency and fairness.

JEL Classification Numbers: C70, D47, D61, D63.
Keywords: Two-sided matching markets, Teacher Assignment, Fairness, Efficiency.

1 Introduction

The reassignment of teachers to schools is a central problem in education policies. We know that teachers are a key determinant of student achievement (Chetty, Friedman, and Rockoff (2014) Rockoff (2004)) so that their distribution across schools can have a major impact on achievement gaps between students from different ethnic or social backgrounds. For instance, most countries' efforts to get more equitable distributions of effective teachers arise from concerns that disadvantaged students may have less access to effective teachers, thereby contributing to sizable achievement gaps.[1] However, any policy playing on the distribution of teachers across schools must be implemented with caution. Indeed, one of the most important issue of the teaching profession is the increasing shortage of qualified teachers (Corcoran, Evans, and Schwab (1994)) and the difficulties to retain new teachers in the profession (Boyd, Lankford, Loeb, and Wyckoff (2005)). An important explanation for both concerns is the inability of the educational systems to accommodate wishes of teachers in terms of their preferred areas to work in. This leaves us with a central question: how can we design an assignment procedure for teachers which would take into account both sides of the market: teachers' preferences as well as schools demands. This paper identifies a unique candidate procedure and empirically assesses its performances.

In many countries, the labor market for teachers is highly regulated by a central administration.[2] The assignment of teachers to schools, for instance, is often managed centraly. Teachers submit ranked lists of wishes to a public administration and each school ranks teachers. The criteria used to rank teachers are diverse, ranging from teachers' performance to standardized test, teachers' experience or geographical distance to a partner.[3] The teacher assignment problem looks quite similar to the college admission problem as defined in Gale and Shapley

[1]Recent initiatives in the U.S. aim at measuring teacher effectiveness and ensuring that disadvantaged students have equal access to effective teachers. These policies (for instance, Race to the Top, the Teacher Incentive Fund, and the flexibility policy for the Elementary and Secondary Education Act) allow states to waive a number of provisions in exchange for a commitment to key reform principles. One could also cite "Teach for America" which recruits and trains teachers who are teaching for at least two years in a low-income community. In the U.K., "Teach First" provides an outstanding training for news teachers.

[2]This is the case in France but also in Italy (Barbieri, Cipollone, and Sestito (2007)), Mexico (Pereyra (2013)) Turkey (Dur and Kesten (2014)) or Uruguay (Vegas, Urquiola, and Cerdàn-Infantes (2006)).

[3]Standardized tests are used, for instance, in Turkey and Mexico while teachers' experience or geographical distance to partners are used, for instance, in France.

(1962). In the latter problem, students have to be assigned to colleges and a *matching mechanism* is a mapping from both students' preferences and schools' rankings over students[4] into possible assignments of students. Typically, one requires three simple desirable properties on this mechanism. Firstly, the assignment produced should be *efficient* in the sense that one cannot find alternative assignments where all students and schools would be weakly better-off and some strictly.[5] Secondly, the assignment should be *fair*, i.e., there should be no student who is refused at a school while other students with lower rankings at that school are accepted (otherwise, we use a standard terminology and say that this student and that school form a *blocking pair*). Finally, one requires *strategy-proofness on the students side*, i.e., that students have incentives to report their preferences truthfully.[6] One of the main result in the literature is that these three properties taken together pin down a unique solution: the deferred acceptance mechanism where students are proposing (Gale and Shapley (1962)) – DA, for short. Thus, any other mechanism will violate one of these three properties and this is certainly an important reason why, in the large majority of cases, this mechanism is used to assign students.[7] Of course, we could also use this mechanism to assign teachers to schools. However, in a first step, our goal in this paper is to point out that this mechanism—or variations on it—will not perform well for the teacher assignment problem. In a second step, we identify a unique (new) mechanism which satisfies a very similar set of desirable properties.

To understand why one cannot simply use the deferred-acceptance mechanism, we note that there is an important difference between the teacher assignment and the college admission problem: when one wants to assign teachers to schools, many of these teachers already have a position and are willing to be re-assigned. In practice, tenured teachers have the right to keep their initial position if they wish, so that the administration has to offer them a position they weakly prefer to the school they are currently assigned to. In other words, the assignment of teachers must be *individually rational*. While fairness has emerged as an important normative criterion in the matching literature, in (the realistic) context where there are tenured teachers willing to be reassigned, all fair matchings may violate the individual rationality constraint.[8] Thus, the prominent mechanism in the college admission problem (DA) fails to be individ-

[4]Criteria used to rank students can also be quite diverse. Typically, they can depend on students' characteristics such as geographic distance to the school, academic performance, social economic status,...

[5]Note that here, one has to adopt an "as if" approach and assume that schools' ranking over teachers can be interpreted as schools' preferences. The basic idea is that these rankings reflect normative criteria which this "as if" approach allows to take into account. A precise discussion of this approach is deferred to the end of the introduction (see "Two-sided efficiency with priorities").

[6]Indeed, most of the time, students are the only strategic entities and schools' orderings are given by law.

[7]It is indeed used in many real assignment problems. For instance, for the assignment of students to high schools in Boston, Hong Kong, New Orleans, New York City,...

[8]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).

ually rational. While no mechanism satisfies the three aforementioned properties together with individual rationality, one can hope to construct mechanisms which accommodate these properties in the best possible way. To achieve this goal, an approach is to use variations on the deferred acceptance mechanism to make it individually rational. For instance, a mechanism identified by the literature and used in practice[9] proceeds as follows: first we artificially modify the school's ordering over teachers so that all teachers initially assigned a school are moved to the top of that school's ranking. Then, once these orderings have been modified, we run DA. While, by construction, this modified version of DA is individually rational, we point out that one looses an important property of DA: it fails to be efficient, i.e., one can reassign teachers in such a way that *both* teachers and schools are better-off (according to the true school's ranking over teachers).[10] In addition, this Pareto-improvement can be done while improving fairness at the same time (i.e., ensuring that the set of blocking pairs shrinks). Our first goal is to build mechanisms which avoid this basic problem: we say that a matching is *two-sided maximal* if (1) it Pareto-dominates the initial assignment, (2) it cannot be improved in terms of (2i) efficiency as well as (2ii) fairness. This requirement is actually quite weak and it is easily shown that two-sided maximal matchings correspond to assignments which are both Pareto-efficient and individually rational on both sides of the market.[11] However, the modification of DA fails to satisfy this property.

In order 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. We can identify these exchanges by finding cycles in an appropriately defined directed graph. The outcome of the BE algorithm depends on the order by which we select the cycles. However, irrespective of the way one selects cycles, we end up with a two-sided maximal matching and, conversely, starting from a two-sided maximal matching, one can find a way to select cycles in the BE algorithm which will eventually yield to this matching. While this result is a useful first step, we end up with a plethora of different possible mechanisms depending on the way we select cycles. However, we show that imposing incentive properties dramatically shrinks the set of possible mechanisms. Our main theoretical result states that there is a unique way to select cycles which makes this algorithm strategy-proof. We provide a mechanism, called the teacher-optimal BE algorithm, which produces two-sided maximal matchings and is the unique strategy-proof mechanism with this property. Interestingly, this

[9] As we will see, this is used in France for the assignment of teacher to secondary public schools. It is also used for the assignment of on-campus housing at MIT, see Guillen and Kesten (2012).

[10] Under (standard) DA, it is well-known that one can reassign teachers and make all of them better-off and some strictly. However, this will be done at the expense of schools given that (standard) DA is in the Core and, hence, efficient. Here, in stark contrast with the standard DA, we show that, under the modified DA, both teachers and schools can be made better-off.

[11] In particular, we can dispense with (2ii) in the definition of two-sided maximality.

mechanism can be characterized using a simple modification of the standard top-trading cycle (Shapley and Scarf (1974)) – TTC, for short: in a first step, one modifies teachers' preferences so that no teacher ranks acceptable a school which finds him unacceptable. Once teachers' preferences have been modified, we run TTC. This result has several implications. First, we see this result as the counter-part, for the teacher assignment problem of the characterization result of the college admission problem. While DA is the unique mechanism which is efficient, fair and strategy-proof in the college admission problem, the teacher-optimal BE algorithm is the unique strategy-proof mechanism which is two-sided maximal (and so cannot be improved in terms of efficiency and fairness). Second, while a priori, we did not want to favor one side of the market and, indeed, the BE algorithm treats schools and teachers symmetrically, once we impose incentive constraints, we end up favoring the teacher side and the only way schools' preferences are taken into account is to ensure that they do not get assigned to an unacceptable teacher. So, imposing strategy-proofness for teachers has an important cost for the school side. This also shows that the teacher assignment problem has a similar "structure" as in the college admission problem. Indeed, in this latter problem, among the set of all fair mechanisms, only the student-optimal is strategy-proof. We show that among the set of two-sided maximal mechanisms, the unique strategy-proof is teacher-optimal.

We also provide additional theoretical results following two directions. Firstly, we consider a case where only teachers are welfare relevant entities. In this context, we provide a similar characterization as the one obtained with the BE algorithm. Here again, we identify a large class of mechanisms. However, while this approach obviously favors teachers, we show – contrary to our main theoretical result – that no mechanism in this class is strategy-proof. Secondly, we simulate matchings in a large market approach where preferences and schools' rankings are drawn randomly from a rich class of distributions.[12] In this context, we show that when the market becomes large, our mechanisms "perform quantitatively better" than the modified DA in terms of utilitarian efficiency and number of blocking pairs. 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 two-sided maximal mechanism. Our arguments build on technique from random graph as in Lee (2014), 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 counterfactuals to quantify the performances of our mechanisms. We confirm that the modified version

[12]These markets can involve quite a large number of agents. For instance, in France, each year, about 65,000 tenured teachers ask for an assignment. In Turkey 8,850 positions were filled by new teachers in 2009 (Dur and Kesten (2014)).

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^*, 30 of them could be simultaneously improved in these two dimensions. Importantly, these 30 fields represent almost all the market we are analyzing since they contain 94.2% 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 dominates that of DA^*. Regarding fairness, 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 in terms of priority of the assigned teachers (compared to the initial assignment) 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.

Two-sided efficiency with priorities. In the main part of this paper, our efficiency notion considers both teachers and schools as welfare-relevant entities (two-sided efficiency). While teachers have real preferences, in the teacher assignment problem, schools' rankings over teachers (hereafter, priorities) are given by law and need not match the true schools' preferences. In the paper we adopt an "as if" approach and do as if schools' priority rankings were schools' preferences and use the two-sided efficiency notion. Our claim is that there is a normative content in schools' priorities and that the "as if" approach allows us to take this into account. To see why, let us fix ideas and consider a typical example where teachers' priorities at schools would be determined by experience, in particular, in disadvantaged schools. The normative content here is clear: it reflects the administration's efforts to assign more experienced teachers to disadvantaged students (see the very first paragraph of the introduction). However, if we were to consider only teachers as welfare-relevant entities, we could end up in situations where teachers could exchange their positions while decreasing the number of experienced teachers in disadvantaged schools. Given the normative content behind priorities, it would be hard to consider this as a Pareto-improvement since disadvantaged students would be hurt by such a re-assignment. Hence, a meaningful requirement is to allow for exchanges of positions across teachers only in cases where this is not at the expense of the experience of teachers in disadvantaged schools. This is exactly what the "as if" approach gives us. In the example we just considered, teachers' priorities are determined by experience. While this is, in virtually all cases, the most important criterion, several other criteria are sometimes considered. For instance, in France, another criteria is "spousal reunification"which gives a bonus of priority to teachers at schools close to where their partner lives. Again, one can easily extract the normative content of these priorities: we do not want to allow reallocation of teachers when this is at the expense of the experience of teachers (in possibly targeted) schools except potentially when it can allow a teacher to join his/her partner. Here again, this is exactly what the "as if" approach will give us.

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 efficiency considerations and fairness 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 fairness 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 fairness"under their weakening. On the contrary, our work keeps the standard definition of blocking pairs and deals with maximal fairness notions using the usual definition. More importantly, our theoretical and empirical results highlight the high cost that maximizing their fairness notion can have in terms of efficiency and in terms of the traditional fairness 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 Appendix F 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.[13] 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. 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 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.

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

In general,

- **Step 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.[14]

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 (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.[15] 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 upon this? The

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

[15]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 $t \succ_s t'$.

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.

Remark 1 *Contrary to what we have in the example, in practice, there are open seats at schools. One may argue that high priority teachers like* t^* *will succeed in getting available seats in schools they desire and that hence the above phenomenon would be very much weakened. However, in an environment where teachers' preferences tend to be similar (i.e., are positively correlated), there will be competition to access good schools. These good schools have a limited number of seats available and one may perfectly imagine that once these open seats are filled by some of the high priority teachers, a similar phenomenon could occur within remaining teachers. We ran simulations in a rich environment (allowing for correlation in teachers'* *and schools' preferences, available seats, new comers and positive assortment in the initial assignment) which confirm this intuition, the results of which are reported in Appendix A.*

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, a 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.[16] 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. 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*.

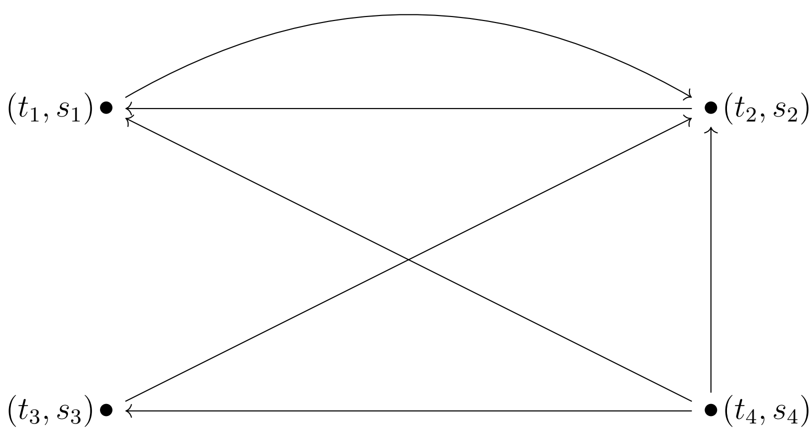
The next example illustrates how the BE algorithm works.

[16] 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.

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

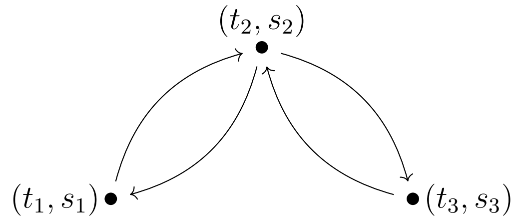
3.1.2 Incentives under Block Exchanges

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.[17] 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 s_i. Teacher t_1's*

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

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[18] 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 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

[18] 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, $\mu_0(t) \in S(k-1)$. As we already noticed, this implies that $\text{Opp}(t, \mu_0, S(k-1))$ is non-empty. Now, because teachers have strict preferences, there is a unique most preferred school for t in $\text{Opp}(t, \mu_0, S(k-1))$.

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 $\text{Opp}(t, \mu_0, S)$) 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, $\varphi(t) \in \text{Opp}(t, \mu_0, S)$. 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 $\text{Opp}(t, \mu_0, S)$ below his initial assignment. By Lemma 2, it implies that at the modified preference profile, TTC is 1-Pareto-dominated by 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 $\text{Opp}(t, \mu_0, S)$ 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. $\text{TO-BE}_t(\succ) \succ_t \varphi_t(\succ) \succ_t \mu_0(t)$. 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 $\text{TO-BE}_t(\succ_t, \succ'_{-t}) \succ_t \varphi_t(\succ_t, \succ'_{-t}) \succ_t \mu_0(t)$.[19] By definition, TO-BE is not affected by t's deviation but then since TO-BE and φ coincide at $\succ'$, we have $\text{TO-BE}(\succ_t, \succ'_{-t}) = \varphi(\succ')$ 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,

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

which contradicts the strategy-proofness of φ.

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.[20] 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.

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

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

t' above t, it does rank first t' *among the individuals who desire* s.[21] 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 $\mu(0) := \mu_0$.

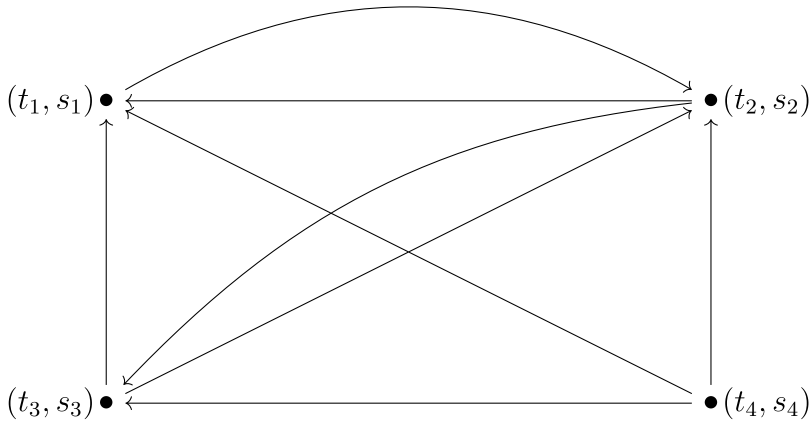
- **Step** $k \geq 1$: Given $\mu(k-1)$, let the teachers and their assignments stand for the vertices of a directed graph where for each pair of nodes (t, s) and (t', s'), there is an edge $(t, s) \longrightarrow (t', s')$ if and only if either (1) teacher t blocks with school s'; or (2) t desires s' 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 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:*

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



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*[22]

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

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

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

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 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 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.[23] In addition, we could assume that school's preferences are drawn in a similar way as students' preferences (allowing tiers), our results would remain essentially the same. That schools' preferences are only based on an idiosyncratic shock is only to simplify the exposition.[24]

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 $\bar{V}_S := 1$ 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 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 occur 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,

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

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

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

[23]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 $\mathbb{R}$.

[24]More precisely, the only issue when introducing a richer class of schools' preferences is that asymptotic stability and individual rationality become incompatible. But if we ignore asymptotic stability, all our results can be extended when allowing the richer class of preferences.

The next three results provide some answers to the three questions stated at the beginning of the current section. The proofs (except for that of Theorem 6) 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 "good" 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 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, using 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 lim inf 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, simulations suggest that a much larger fraction of teachers are not moving. In addition, these simulations show that the assumption we made above that $u_1 > u_2 + 1$ is not necessary and that the result seems to hold in much broader contexts.[25] 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 that 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, the intuition can be seen as follows: for a given tier k, for any school in S_k and any agent initially matched to such a school in S_k, the probability that they both enjoy high idiosyncratic payoffs for each other is bounded away from 0. Thus, as the market grows, with probability approaching one, one can appeal to the Erdös-Rényi's result on the existence of a (perfect) matching within the set of individuals and schools. However, there are two difficulties here. First, one has to ensure individual rationality of the

[25] Available upon request.

matching so obtained which we handle by restricting the set of teachers and schools to those who have idiosyncratic payoffs for their initial match bounded away from the upper bound. Second, in the sketched we just provided we are implicitly assuming that the designer has access to agent's cardinal utilities. However, in this paper, we assume—as usually the case in practice—that matching mechanisms map ordinal preferences into matchings and a large part of the proof is devoted to handle this issue.

Now, as we already pointed out, while the BE algorithm treats symmetrically teachers and school, TO-BE favors teachers at the expense of schools. Thus, it is natural to expect that TO-BE is asymptotically teacher efficient. In addition, TO-BE only ensures that schools get assigned a teacher that they weakly prefer to their initial assignment. Hence, for each school, its assignment under TO-BE is a random draw within the set of teacher that it finds acceptable. Thus, given its idiosyncratic payoff for its initial assignment $\eta_{\mu_0(s)s}$, the expected payoff of a school s for TO-BE (s) is $\mathbb{E}\left[\eta_{st}\left|\eta_{st} \geq \eta_{\mu_0(s)s}\right.\right] = \frac{1}{2}\left(1 + \eta_{\mu_0(s)s}\right)$. Thus, the (unconditional) expected payoff of school s for TO-BE (s) is $\mathbb{E}\left[\frac{1}{2}\left(1 + \eta_{\mu_0(s)s}\right)\right] = \frac{3}{4}$. Thus, TO-BE cannot be asymptotically school efficient.

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, a 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 that 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 with a brief description of the French procedure. We also provide a short presentation of the data set we use. 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.[26] 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.[27]
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). Newly tenured teachers and teachers who want to move to another region submit an ordered list of regions. A matching mechanism (described below) is used to match teachers to regions using priorities defined by the point system. This phase is managed by the central administration.

[26] 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/

[27] 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

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 the 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.

Each teacher only teaches one subject. Hence, the assignment process is decomposed into several markets—one for each subject taught—so that positions are specific to a subject and markets can be considered as being independent from one subject to the other.[28] In each market, the mechanism used in both phases is working as follows: 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 yields to the outcome of the *Teacher Proposing Deferred Acceptance* according to the modified priorities.[29] Hence, the mechanism used to assign French teachers to public schools is equivalent to DA* - as defined in Section 2.

In 2013, just over 25,000 teachers applied in Phase 1, and around 65,000 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 2,500 teachers), Contemporary Literature (around 2,000 teachers), Mathematics (around 2,000 teachers), while others are smaller like Thermal Engineering (around 60 teachers) or Beauticians (around 15 teachers) with a wide range of subjects in between.

[28]In practice, couples from different fields can submit joint applications which ties up the fields. However, we eliminated all couples from our sample. Details are provided in Appendix G.

[29]In this paper we ignore the issues with coarse priorities. However, in the French system, teachers' priorities at schools can be coarse. Hence, in practice, the algorithm starts by breaking ties (using teachers' birth dates). Once ties are broken, School Proposing Deferred Acceptance is run using the modified priorities *with no ties* and the reported preferences. From this outcome Stable Improvement Cycles are run using again the modified (strict) priorities. Thus, the outcome is equivalent to the Teacher Proposing Deferred Acceptance which in turn may be Pareto-dominated by a Teacher-Optimal Stable mechanism. Our mechanisms and results can be easily extended to an environment with coarse priorities.

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 the 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 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. However, there is no such binding constraint in the first phase: the assignment of teachers to the 31 regions. Since the mechanism at use is DA*, it is a dominant strategy for all agents to be truthful. Thus, we take for granted that agents' reported preferences in Phase 1 are the true preferences in order to run our counterfactuals. In addition, given the agents' assessments over schools they may obtain in the second phase, agents have well defined preferences over 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 are 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. We end this section by providing two pieces of information on this market.

Fact 0 *(i) Under the regular DA mechanism, there are at least 1325 teachers who see their individual rationality violated, i.e., who get assigned a region that they consider worse than their initial region.*[30] *(ii) The individually rational mechanism which maximizes movement allows 2,257 teachers to move from their initial assignment.*[31]

This fact shows that the regular DA mechanism is indeed not individually rational in this

[30] In our data set, for each teacher, we have the reported preferences only up to his initial region. Hence, we do not know how teachers rank regions below their initial assignment. However, one can show that when running DA on these truncated preferences, the number of unassigned teachers is a lower bound on the number of teachers who see their individual rationality violated when running DA on the full preference lists.

[31] In order to find such an assignment, we build a bipartite graph with teachers on one side and schools on the other side. We consider the complete bipartite graph where each edge will be associated with a weight. We put weight ∞ on edges (t, s) where s is unacceptable to t (i.e., worse than his initial assignment). We put weight 1 on the edge if t is initially matched to s. Finally, we put weight 0 on all other edges (i.e., if t finds s strictly better than his initial assignment). The weight of a matching is defined as the sum of weights over all his edges. We used a standard algorithm to find a matching with minimal weight (see Kuhn (1955) and Munkres (1957)). It is easily checked that such a matching maximizes movement among all individually rational matchings.

market and the violation of individual rationality is quite strong. The second point gives us a sense in which there is congestion on this market: if we only focus on individually rational matchings and try to ensure as much movement as possible, 21% of teachers will be able to move. We should keep in mind this upper bound when considering the performances of our algorithms and the scope of their improvements.[32]

4.3 Results

4.3.1 Preliminaries: many-to-one

Before moving to the description of the results, we need to discuss briefly the generalization of our mechanisms to the many-to-one framework. A school/region can now be assigned several teachers and, starting with preferences over single teachers, we need to define schools' preferences over bundle of teachers. We adopt a very conservative approach here which will only strengthen our main empirical findings. Consider a school/region with q positions to fill and two vectors of size q, say $\mathbf{x} := (t_1, \ldots, t_q)$ and $\mathbf{y} := (t_1, \ldots, t_q)$. Let us assume that each of these vectors are ordered in such a way that for each $k = 1, \ldots q-1$, the kth element of vector $\mathbf{x}$ is preferred to its $k+1$th element; and similarly for vector $\mathbf{y}$. We say that $\mathbf{x}$ is preferred by the school/region to $\mathbf{y}$ if for each $k = 1, \ldots q$, the kth element of vector $\mathbf{x}$ is (weakly) preferred to the kth element of vector $\mathbf{y}$.

With this definition in mind, all our concepts (two-sided maximality or one-sided maximality) can be naturally extended. Mechanisms characterizing these notions can also be easily extended. Because, these are the mechanisms we used to run our counter-factuals, we state them precisely in Appendix F.

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 outcomes within this class. As is made clear in Appendix F, in a many-to-one environment, TO-BE is extended to a class of mechanisms parametrized by an ordering over teachers. We randomly select an ordering and hence randomly select an outcome in this class. We randomly draw selections for each mechanism ten times. In addition, there are indifferences in priorities of regions over teachers.[33] We use a single tie-breaking rule to break ties in regions' priorities. We draw

[32] The relatively small fraction of teachers being able to 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 only ask for one region (beyond their initial region). Combined with correlation in preferences, this structurally restricts the possibility of movement in the market.

[33] Many young teachers use only one criteria - number of years of experience- to compute their priorities, so that they have the same priority in a given region.

randomly ten times the tie-breaking rule. In the end, the results reported in Tables 1 to 3 for BE, TO-BE and 1S-BE correspond to average outcomes over one hundred draws – ten random selection of tie-breaking rules times ten random selection of outcomes. The results for DA* correspond to an average over ten iterations of tie-break only. The results for the BE algorithm and the 1S-BE are successively presented in the next section.

4.3.2 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. In practice, a first 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. This means that DA* is not two-sided efficient which is a necessary condition for two-sided maximality. Fact 1 below illustrates that point in our data:

Fact 1 *DA* is not two-sided maximal in 33 subjects out of 49. These subjects represent 95.9% of the teachers.*[34]

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. Regarding teachers' welfare, two results reported in Table 1 are worth commenting. First, on average, BE$\circ$DA* more than doubles the number of teachers who obtain a new region compared to DA*: 565 teachers are moving from their initial allocation under DA* versus 1487.6 under BE$\circ$DA*. Second, the same table reports the cumulative distribution of the number of teachers who obtain school rank k. While we know from the theory that the distribution of the rank obtained by teachers under BE$\circ$DA* first order stochastically dominates this same distribution under DA*, the dominance is indeed significant.

Regarding regions' welfare, there are several possible measures. We focus below on one natural approach; however, we tested the robustness of our results using alternative approaches

[34] Even if BE$\circ$DA* = DA*, it could be the case that DA* is not two-sided maximal. Indeed, in our definition of two-sided maximality, we impose that schools must not be hurt compared to the initial allocation. Since DA* may hurt schools, it can be two-sided Pareto efficient but still violate two-sided maximality. In 30 subjects, BE$\circ$DA* $\neq$ DA* and in 3 subjects BE$\circ$DA* = DA* but DA* hurts the welfare of at least one region compared to the initial assignment. Finally, we note that if we restrict our attention to the 19 subjects with more than 100 teachers, in only one subject DA* is two sided maximal.

Table 1: Welfare of teachers under different mechanisms

Choice	Init	DA*	TO-BE	BE(Init)	BE(DA*)	1S-BE(Init)	1S-BE(DA*)
1	0	422.5	1158.6	1152.6	1169.2	1355.4	1347.2
2	7935	8084.3	8314.5	8286.8	8290.2	8385.8	8371.0
3	9125	9220.6	9359.0	9336.5	9341.9	9399.2	9390.8
4	9743	9796.1	9900.2	9882.7	9884.9	9929.1	9917.7
5	10038	10077.8	10148.0	10137.2	10140.2	10170.5	10162.8
6	10271	10297.0	10339.5	10328.4	10331.0	10354.9	10351.2
7	10366	10383.5	10417.5	10409.2	10408.7	10426.5	10422.4
8	10420	10432.5	10457.7	10450.2	10450.6	10463.7	10461.0
9	10461	10474.5	10492.2	10485.5	10487.5	10495.5	10494.6
>= 10	10579	10579	10579	10579	10579	10579	10579
Nb teachers moving	0	565.0	1363.4	1460.6	1487.6	1732.0	1708.7
Min	0	560.0	1334.0	1416.0	1456.0	1696.0	1684.0
Max	0	568.0	1401.0	1517.0	1513.0	1768.0	1739.0

† Notes: This table presents the cumulative distribution of the number of teachers who obtain school rank k under their initial assignment in column 1, under DA* in column 2, TO-BE in column 3, BE(Init) in column 4, BE(DA*) in column 5, 1S-BE(Init) in column 6 and 1S-BE(DA*) in column 7. Data from the assignment process of French teachers to regions in 2013.

which yield no significant differences in the results. Given a mechanism, we look at the improvement a region obtains (from the initial assignment) in terms of number of positions improved. More precisely, given a region, we first take the initial assignment and sort it by decreasing order of priorities. We obtain a vector where the first element/position is the teacher with the highest priority in that region at the initial assignment, the second element/position is the teacher with the second highest etc... Call this vector $\mathbf{x}$. We do the same operation for the assignment of this very same region but now with the mechanism under study. Let us call this vector $\mathbf{y}$. Finally, we say that a position k is assigned a teacher with higher (resp., lower) priority if the kth element of vector $\mathbf{y}$ has a higher (resp., lower) priority than the kth element of vector $\mathbf{y}$. Based on this, we compute the percentage of net improvement of positions, i.e., the percentage of positions receiving a teacher with a higher priority, minus the percentage of positions being assigned a teacher with a lower priority.

Table 2 reports, for the different mechanisms we run, the cumulative distribution of the percentage of net improvement of positions, i.e., for each percentage x, the proportion of regions having less than x percent of net improvement of positions. Again, we observe that the distribution under BE∘DA* first order stochastically dominates this same distribution under DA*.

Finally, we compare the performance of DA* and BE∘DA* in terms of fairness. The first row of Table 3 reports that on average 2496.5 teachers are not blocking under DA* and 3798.3 under BE$\circ$DA*, which represents an increase of 52.1 % of the number of teachers who are not blocking with any region. More generally, we observe that fairness is significantly increased.[35]

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 assignment achieved by 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 for an arbitrary outcome of the BE mechanism its set of blocking pairs may differ from that of DA* and, similarly, the outcome may not 2-Pareto-dominate DA* . 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 – which are two-sided maximal – can be expected to perform much better. Second, our theoretical results (Theorems 6, 7 and 8) suggest that BE and TO-BE perform better in large markets than DA*.

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* : on average 565 teachers obtain a new assignment under DA*, versus 1460.6 under BE and 1363.4 under TO-BE.

Fact 2 *The distribution of ranks obtained by teachers under TO-BE first order stochastically dominates the distribution of BE, which dominates the one of DA*.*

Note however, that there is no 2-Pareto domination between the matchings: some teachers may prefer their assignment under DA* to the one they obtain under BE or TO-BE.[36]

[35] 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 imposing the individual rationality constraint.

[36] On average, 258.5 teachers obtain a region they rank strictly higher under DA* than under BE (254.5 under TO-BE). Conversely, 1152.7 teachers strictly prefer their assignment under BE than under DA*, and 1096.7 under TO-BE.

Table 2: Welfare of regions under different mechanisms

Net percentage of positions	DA^*	TO-BE	BE(Init)	BE(DA^*)	1S-BE(Init)	1S-BE(DA^*)
-100/-91%	0.18	0	0	0.18	0.79	0.71
-90/-71%	0.18	0	0	0.18	1.06	0.93
-70/-51%	0.31	0	0	0.31	1.75	1.39
-50/-31%	0.58	0	0	0.50	3.21	2.55
-30/-1%	1.04	0	0	0.93	4.88	3.99
0%	84.32	71.96	71.17	70.92	72.71	71.82
1/29%	87.94	75.58	74.58	74.74	76.58	76.05
30/49%	91.05	79.15	77.99	77.67	80.03	79.17
50/69%	94.83	85.05	84.31	84.02	85.77	84.94
70/89%	97.40	90.17	89.69	89.02	89.66	88.79
90/100%	100	100	100	100	100	100
% of regions with no priority change	83.21	71.96	71.17	69.84	67.07	67.14

† Note: this table presents the cumulative percentage of regions having a net welfare improvement (compared to their initial assignment). 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^* 0.18 % of the regions have between 91% and 100 % of their seats which are assigned a teacher with a lower priority (in net).

Regarding stability, BE and TO-BE also perform significantly better than DA^*. Table 3 shows that the average number of teachers not being part of a blocking pair increases from 2496.5 under DA^* to respectively 3729.3 and 3746.9 under BE and TO-BE.

Fact 3 *The three distributions of the number of regions with which teachers can block under DA^*, BE and TO-BE can be ranked stochastically: DA^* is dominated by BE which is dominated by TO-BE.*[37]

Finally, comparing regions' welfare across mechanisms is of particular interest as we know that DA^* can hurt some regions, contrary to BE and TO-BE. This is confirmed by Table 2. Under DA^*, 1.04 % of the regions have at least 1% of their positions which are assigned a teacher with a lower priority than under the initial assignment. On the contrary, under BE and TO-BE, no region has a position where the teacher assigned to it has a lower priority than the teacher initially assigned to that position.

[37] 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^*.

Fact 4 *The distributions of the percentage of net improvement of positions can be stochastically ordered : the distribution of DA* is dominated by the one of TO-BE, which is dominated by the one of BE.*

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 (Theorem 7 and 8).

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*.[38] The next section tests if we could further improve upon DA* by relaxing the constraint that no region should be hurt (compared to the initial assignment). In order to do so, we provide empirical evidence on the performance of 1S-BE, the one-sided maximal algorithm we have defined in Section 3.2.

Table 3: Stability of the matching obtained with different mechanisms

Nb regions	Init	DA*	TO-BE	BE(Init)	BE(DA*)	1S-BE(Init)	1S-BE(DA*)
0	1980	2496.5	3746.9	3729.3	3798.3	3940.9	4005.9
1	8722	8881.0	9250.9	9214.6	9235.9	3293.9	9307.8
2	9694	9787.9	10004.3	9983.1	9992.3	10026.8	10036.0
3	10096	10149.5	10299.3	10287.6	10293.4	10309.8	10312.7
4	10323	10360.3	10447.4	10438.7	10444.9	10457.9	10459.3
>= 5	10579	10579	10579	10579	10579	10579	10579
Nb of teachers blocking with at least one region							
Mean	.	8082.5	6832.1	6849.7	6780.7	6638.1	6573.1
Min	.	8078.0	6728.0	6703.0	6677.0	6527.0	6453.0
Max	.	8087.0	6933.0	6985.0	6917.0	6809.0	6685.0

† Notes: The upper part of this table presents the cumulative distribution of the number of regions teachers are blocking with. Data from the assignment process of French teachers to regions in 2013. Column 1 reports the cumulative distribution of the number of regions teachers are blocking with under their initial assignment. The following columns report the cumulative distribution of the number of regions teachers are blocking with under respectively, DA*, TO-BE, BE(Init), BE(DA*), 1S-BE(Init) and 1S-BE(DA*).

[38]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.

4.3.3 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. 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 three the number of teachers moving compared to DA* and increases it by 15% 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 31 subjects. 95.3% of the teachers belong to these subjects. In one subject, DA* is two sided maximal but not one sided maximal.*

We now turn to the results on 1S-BE starting from the initial allocation (referred to as 1S-BE(Init) in our tables and graphs) to compare its performance with the other mechanisms. Regarding teachers' welfare and fairness, the 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*.[39]

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

Fact 6 *Under 1S-BE, in 4.9 % of the regions, the percentage of positions filled by a teacher with a lower priority is higher than the percentage of positions filled by a teacher with a higher priority. This is in contrast with BE or TO-BE under which, by definition, regions cannot be assigned teachers with a lower priority compared to the initial assignment.*

References

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

Barbieri, G., P. Cipollone, and P. Sestito (2007): "Labour Market for Teachers: Demographic Characteristics and Allocative Mechanisms," *Giornale degli Economisti e Annali di Economia*, 66, 335–373.

[39] 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.

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff (2005): "Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students," *American Economic Review*.

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.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014): "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*.

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

Corcoran, S., W. Evans, and R. Schwab (1994): "Changing labor-market opportunities for women and the quality of teachers, 1957–2000," *American Economic Review*.

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.

Grimmett, G., and C. McDiarmid (1975): "On colouring random graphs," *Mathematical Proceedings of the Cambridge Philosophical Society*, 77, 313–324.

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.

Kuhn, H. W. (1955): "The Hungarian Method for the Assignment Problem," *Naval Research Logistics Quarterly*, 2(1-2), 83–97.

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.

Munkres, J. (1957): "Algorithms for the Assignment and Transportation Problems," *Journal of the Society for Industrial and Applied Mathematics*, 5(1), 32–38.

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

In this section we report a certain number of simulations. Our goal here is to argue that in a realistic setting where a significant fraction of teachers are new comers and where each school has a significant fraction of available seats, the effect identified in Example 1 can occur "quite often". In order to do so, we consider the following environment. We have 600 teachers and 30 schools with 20 seats each. 300 teachers are new comers. The other half of teachers initially have an assignment. We assume that the total number of seats equals the total number of teachers so that half of the seats of each school are occupied and the other half are open seats.

We randomly draw the utility of teacher t for school s as follows:

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

where u_s is the common value of school s drawn uniformly over the interval $[0, a]$ and ξ_{st} is the idiosyncratic shock drawn uniformly over the interval $[0, 1]$. Parameter a measures the degree of correlation in teachers' preferences. For schools' preferences, we have a similar specification:

$$V_s(t) = v_t + \eta_{st}.$$

We assume that both common values (v_t) and idiosyncratic shocks (η_{st}) are drawn uniformly in $[0, 1]$.

Finally, given the dynamic nature of the teacher assignment process, teachers initially assigned good schools have on average a higher priority than teachers assigned less good schools. In order to take this into account, we assume that the initial assignment is assortative: teachers with high common values are initially matched to schools with high common values. Formally, we assume that the 10 teachers t having a common value v_t within the 10 highest are matched to the school s with the highest common value u_s. Similarly, the next 10 teachers having a common value within the next 10 highest are matched to the school with the second highest common value, and so on.

We draw 50 times teachers' and schools' preferences. For each draw we computed the outcome of DA* as well as an outcome on the Pareto frontier which Pareto-dominates DA*.[40] The following table reports the number (averaged across iterations) of teachers staying at their initial assignment for both DA* and the Pareto-dominating matching for several possible values of the correlation parameter a.[41]

[40] More precisely, in order to find the Pareto-dominating outcome we ran TO-BE starting from the allocation given by DA*. See Section 3.1 for the definition of TO-BE.

[41] In our environment, teachers always prefer to be matched rather than being unmatched. Given that the total number of seats equals the number of teachers, all teachers who are initially unmatched will end up being matched under DA* and so under the Pareto-dominating matching.

Table 4: Simulations results

[0,a]	DA* init	Pareto-dominating matching
[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

In many instances DA* can be improved both in terms of efficiency and stability as in Example 1. The intuition is essentially the same: within top schools (i.e., for instance, the schools corresponding to the ten highest common values), open seats are filled very quickly by teachers initially matched to these top schools but also partly by teachers initially matched to less good schools. Hence, at some point, we are back to the environment of the example where teachers initially matched to good schools are willing to move to other good schools but cannot do so because of the large set of remaining tenured teachers – initially matched to less good schools – among whom some teacher may have a higher priority at the school they are targeting.

The simulations reveal that the scope of the improvement can potentially be large depending on the precise value of the correlation parameter a (it is maximized for intermediate values of the parameter).

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

In the sequel, we prove our characterization result of one-sided maximal matchings given in Theorem 4. Our proof is divided into two parts. We start by showing that any outcome of the 1S-BE algorithm is a one-sided maximal matching (Section C.1):

Proposition 7 *If μ is an outcome of the 1S-BE algorithm then μ is one-sided maximal.*

Then, we move to the proof that any one-sided maximal matching corresponds to a possible outcome of the 1S-BE algorithm (Section C.2):

Proposition 8 *If μ is one-sided maximal then μ is an outcome of the 1S-BE algorithm.*

C.1 Proof of Proposition 7

Before moving to the proof we introduce a new notation. Given matching μ, we denote $\mathcal{B}_\mu$ for the set of blocking pairs of μ.

In the sequel, we fix two matchings μ and μ' such that μ' Pareto-dominates μ for teachers and $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_\mu$. We show below that starting from μ, the graph associated to the 1S-BE algorithm must have a cycle. Hence, any outcome of 1S-BE must be one-sided maximal, as claimed in Proposition 7.

To give the intuition of each step of the proof, which uses a lot of graphical arguments, we will use an example to illustrate each part. This example involves 6 teachers, $t_1, t_2, t_3, t_4, t_5, t_6$ and 6 schools $s_1, s_2, s_3, s_4, s_5, s_6$. In the example, matchings μ and μ' are as follows:

$$\mu = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 & t_5 & t_6 \\ s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \end{pmatrix}$$
$$\mu' = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 & t_5 & t_6 \\ s_2 & s_3 & s_4 & s_5 & s_6 & s_1 \end{pmatrix}$$

As in Lemma 1, we can exhibit "cycles of exchanges" which can be used to go from μ to μ' in the proposition.

In Lemma 1, these cycles of exchanges were actual cycles in the graph associated with BE. However, when considering the graph associated with 1S-BE, this is not the case anymore: the cycles of exchanges are not necessarily cycles of the graph associated with 1S-BE. Before moving to the first lemma, we note that all the nodes that are not part of cycles of exchanges are those where the teacher of that node has the same allocation between μ and μ'. In the following the "nodes of the cycles of exchanges" will be all the nodes (t, s) s.t $\mu(t) \neq \mu'(t)$. We will say that a node (t, s) 1S-BE-points to another node (t', s') if (t, s) points toward (t', s') in the graph associated with the 1S-BE algorithm (starting from μ).

Lemma 6 *Fix a node (t, s) of the cycles of exchanges. Then:*

1. *either its predecessor according to the cycles of exchanges 1S-BE-points toward (t, s);*
2. *or there is a node (t', s') in the cycles of exchanges such that t' does not block with s under μ, $s \succ_{t'} s'$ and t' has the highest priority among those who desire s but do not block with it under μ. And so (t', s') 1S-BE-points toward (t, s).*

Before moving to the proof, let us illustrate this lemma in the example. Assume that all the nodes except (t_3, s_3) are 1S-BE-pointed by their predecessors in the cycle of exchanges. According to Lemma 6 there must be a node (t', s') in the cycle of exchanges which 1S-BE-points toward (t_3, s_3). In the graph of Figure 1, this node is assumed to be (t_5, s_5). The dashed edge from (t_2, s_2) to (t_3, s_3) is here to show that this is not an edge of the 1S-BE graph but it is only an edge corresponding to the cycle of exchanges.

PROOF. Call (t'', s'') the predecessor of node (t, s) in the cycles of exchanges so that $s'' := \mu(t'')$ and $s := \mu'(t'')$. Because μ' Pareto-dominates μ for teachers, we know that $s \succ_{t''} s''$ so that t'' desires s under μ. Assume that (t'', s'') does not 1S-BE-point to (t, s). This means that t'' does not block with s under μ and that there is another teacher t' who does not block with s and has the highest priority among those who desire s and do not block with it. Thus, (t', s') (where $s' := \mu(t')$) 1S-BE points toward (t, s). It remains to show that (t', s') is part of the cycles of exchanges. If it was not the case, it would mean that $\mu(t') = \mu'(t') = s'$. Let us recap. We have that t' does not block with s under μ. In addition, by definition of t', we must have that $t' \succ_s t''$ (since t'' does not block with s under μ and desires s). In addition, t' desires s under μ and so $\mu(t') = \mu'(t')$ implies that t' also desires s under μ'. Hence, because $t'' \in \mu'(s)$, we obtain that t' blocks with s under μ'. This contradicts our assumption that $\mathcal{B}_{\mu'} = \mathcal{B}_{\mu}$. ∎

Lemma 6 allows us to identify a subgraph (N', E') of the 1S-BE graph starting from μ such that N' are the nodes of the cycles of exchanges and the set of edges E' is built as follows. We start from $E' = \emptyset$ and we add the following edges: for each node (t, s) in the cycles of exchange, if its predecessor $(\tilde{t}, \tilde{s})$ under the cycles of exchanges 1S-BE-points to (t, s) then $((\tilde{t}, \tilde{s}), (t, s))$ is added to E'. If on the contrary, $(\tilde{t}, \tilde{s})$ does no not 1S-BE-points to (t, s), then we pick the node (t', s') in the cycles of exchanges identified in the second condition of Lemma 6 which 1S-BE-points toward (t, s) and we add $((t', s'), (t, s))$ to E'. Note that, by construction, each node in N' has a unique in-going edge in (N', E'). In the example, this subgraph (N', E') is given by the right graph of Figure 1 (the solid arrows). Note that this graph admits a cycle: $(t_3, s_3) \to (t_4, s_4) \to (t_5, s_5) \to (t_3, s_3)$. This is a simple property of digraphs with in-degree one:

Lemma 7 *Fix a finite digraph (N, E) such that each node has in-degree one. There exists a cycle in this graph.*

Proof. Fix a node n_1 in the graph (N, E). Because it has in-degree one, we can let n_2 be the unique node pointing to n_1. Again from n_2 we can let n_3 be the unique node pointing to n_2. Because there is a finite number of nodes in the graph, this process must cycle at some point. ∎

As the example illustrates, applying this lemma to (N', E') leads to the following corollary:

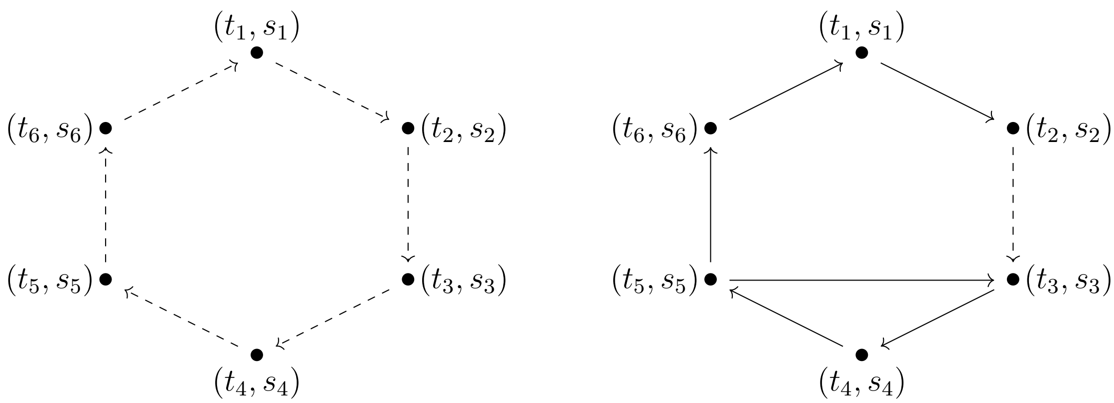
Corollary 3 *There is a cycle in the subgraph (N', E').*

We are now in a position to prove Proposition 7.

Completion of the proof of Proposition 7. Let μ be an outcome of the 1S-BE algorithm. Proceed by contradiction and assume that μ is not one-sided maximal. Thus, there must exist a matching μ' such that μ' Pareto-dominates μ for teachers and $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\mu}$. Corollary 3 implies that there must be a cycle in the graph associated with 1S-BE starting from μ, contradicting the fact that μ is an outcome of 1S-BE. ∎

Figure 1: Cycles of exchanges and (N', E').

a) Cycle of exhanges b) (N', E')



C.2 Proof of Proposition 8

In the sequel, we fix a one sided maximal matching μ'. We let μ be a matching such that μ' Pareto-dominates for teachers μ and satisfies $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\mu}$. We claim that there is a cycle in the graph associated with 1S-BE starting from μ which, once implemented, leads to a matching $\tilde{\mu}$ such that μ' Pareto-dominates $\tilde{\mu}$ for teachers and satisfies $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\tilde{\mu}}$. Note that this implies Proposition 8. Indeed, because, by definition, μ' Pareto-dominates μ_0 and $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\mu_0}$, we must have a cycle in the graph associated with 1S-BE starting from μ_0, which once implemented, yields to a matching say $\tilde{\mu}_1$ such that μ' Pareto-dominates $\tilde{\mu}_1$ for teachers and satisfies $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\tilde{\mu}_1}$. Now, we can iterate the reasoning and we get again that there is a cycle in the graph associated with 1S-BE starting from $\tilde{\mu}_1$, which, once implemented, yields to a matching say $\tilde{\mu}_2$ such that μ' Pareto-dominates $\tilde{\mu}_2$ for teachers and satisfies $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\tilde{\mu}_2}$. We can pursue this reasoning. At some point, because the environment is finite, we must reach matching μ', as was to be shown.

In the sequel, as in the proof of Proposition 7, we consider the digraph (N', E') as built in Section C after Lemma 6. Consider a cycle $\tilde{C}$ in this graph (which exists by Lemma 7). Let $\tilde{\mu}$ be the matching obtained once the cycle $\tilde{C}$ is implemented. In the example introduced in Section C, this matching would be:

$$\mu_1 = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 & t_5 & t_6 \\ s_1 & s_2 & s_4 & s_5 & s_3 & s_6 \end{pmatrix}$$

We first show the following lemma:

Lemma 8 *μ' Pareto-dominates $\tilde{\mu}$ for teachers.*

PROOF. Fix a teacher t. If the node (t, s) to which t belongs is not part of the cycles of exchanges, we know t does not move from μ to μ' and so (t, s) is not in the cycle $\tilde{C}$. Hence, $\mu(t) = \tilde{\mu}(t) = \mu'(t)$. So assume that (t, s) is part of the cycles of exchanges and let $s := \mu(t)$ and $s' := \mu'(t)$ with $s \neq s'$. There are three possible cases:

- *Case 1:* $s = \tilde{\mu}(t) \neq s'$. Because μ' Pareto-dominates μ for teachers, we have that $\mu'(t) = s' \succeq_t \tilde{\mu}(t) = \mu(t) = s$.

- *Case 2:* $s \neq \tilde{\mu}(t) = s'$. In such a case, we trivially have $\mu'(t) \succeq_t \tilde{\mu}(t)$.

- *Case 3:* $s \neq \tilde{\mu}(t) := s_1 \neq s'$. By construction of the graph (N', E') when we implement cycle $\tilde{C}$, we know that there is a unique edge $\big((t, s), (t_1, s_1)\big)$ in $\tilde{C}$ and that (t, s) is not the predecessor of (t_1, s_1) under the cycles of exchanges, since otherwise, t would be matched to s' under $\tilde{\mu}$ which is not the case by assumption. Hence, by construction of (N', E'), the predecessor of (t_1, s_1) under the cycles of exchanges, say (t'', s''), does not 1S-BE point to (t_1, s_1) and, in addition, t does not block with s_1 under μ, $s_1 \succ_t s$ and t has the highest priority among those who desire s_1 but do not block with it under μ and 1S-BE-points to (t_1, s_1). Because (t'', s'') does not 1S-BE point to (t_1, s_1), we know that t'' does not block with s_1. While because (t'', s'') points to (t_1, s_1) under the cycles of exchange, we must have that t'' desires s_1. Thus, we conclude that $t \succ_{s_1} t''$.

  Now, proceed by contradiction and assume that $(\tilde{\mu}(t) =) s_1 \succ_t s' (= \mu'(t))$. Because $t'' \in \mu'(s_1)$ (recall that (t'', s'') is the predecessor of (t_1, s_1) under the cycles of exchange) and $t \succ_{s_1} t''$, we have that t blocks with s_1 under μ' i.e. $(t, s_1) \in \mathcal{B}_{\mu'}$. But, as already claimed, $(t, s_1) \notin \mathcal{B}_{\mu}$. This contradicts that $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\mu}$. Thus, we must have $\mu'(t) \succeq_t \tilde{\mu}(t)$.[42]

So we have shown that $\forall t$, $\mu'(t) \succeq_t \tilde{\mu}(t)$. ∎

The following lemma completes the argument.[43]

---

[42]Case 3 in Lemma 8 can be illustrated in the example. The node (t, s) would be (t_5, s_5) in the right graph of Figure 1. t_5 is matched to s_3 under $\tilde{\mu}$ but is matched to s_6 under μ'. Under $\tilde{C}$ (i.e., $(t_3, s_3) \to (t_4, s_4) \to (t_5, s_5) \to (t_3, s_3)$), node (t_5, s_5) points to (t_3, s_3) while (t_2, s_2) does not 1S-BE-point to (t_3, s_3). Because (t_2, s_2) points to (t_3, s_3) in the cycle of exchanges, it means that $t_2 \in \mu'(s_3)$ so that if t_5 preferred s_3 to his match under μ', s_2, it would imply that t_5 blocks with s_3 under μ' while he does not under μ and so this would yield the contradiction.

[43]Note that up to here, all the arguments we provided can be applied to the many-to-one environment. However, the following Lemma explicitly uses the one-to-one environment and is not true anymore in many-to-one. However, we can use additional arguments to show that Proposition 8 goes through in the many-to-one setting. This is provided in Section S.2 of the supplementary material.

Lemma 9 *We have that $\mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\tilde{\mu}}$*

Proof. The proof proceeds by contradiction. Assume that there is a teacher t and a school s s.t $(t,s) \in \mathcal{B}_{\mu'}$ but $(t,s) \notin \mathcal{B}_{\tilde{\mu}}$. Note first that teacher t desires s under $\tilde{\mu}$ because $s \succ_t \mu'(t)$ (by $(t,s) \in \mathcal{B}_{\mu'}$) and $\mu'(t) \succeq_t \tilde{\mu}(t)$ (by Lemma 8). So because $(t,s) \notin \mathcal{B}_{\tilde{\mu}}$ we must have that $\tilde{t} \succ_s t$ for $\tilde{t} := \tilde{\mu}(s)$. Since $(t,s) \in \mathcal{B}_{\mu'} \subseteq \mathcal{B}_{\mu}$, we know that t blocks with s under μ and so, since we are in a one-to-one setting, $\mu(\tilde{t}) \neq s$. This implies that $(\tilde{t}, \mu(\tilde{t}))$ is part of cycle $\tilde{C}$. Since $(t,s) \in \mathcal{B}_{\mu'}$, we also know that $\mu'(\tilde{t}) \neq s$. So to recap, we have that $\mu(\tilde{t}) \neq \tilde{\mu}(\tilde{t}) \neq \mu'(\tilde{t})$. But, this means that in the graph (N', E'), $(\tilde{t}, \mu(\tilde{t}))$ points to $(s, \mu(s))$ while $(\tilde{t}, \mu(\tilde{t}))$ is not the predecessor of $(s, \mu(s))$ in the graph of exchanges. By construction of (N', E') this means that $\tilde{t}$ does not block with s under μ and has the highest priority among teachers who desire s under μ and do not block with it under μ. In particular, because $\tilde{t}$ does not block with s under μ (but desires it under μ) we must have $\mu(s) \succ_s \tilde{t}$. In addition, since $(t,s) \in \mathcal{B}_{\mu}$, we must have $t \succ_s \mu(s) \succ_s \tilde{t}$, contradicting that $\tilde{t} \succ_s t$. ∎

D Proof of Theorem 5

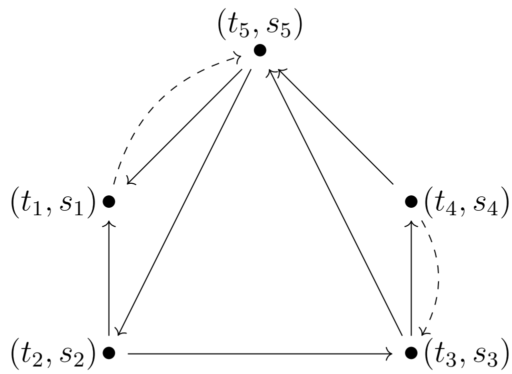
In order to prove this result, we exhibit an instance where, irrespective of which (sequence of) cycle(s) one selects in the graphs associated with 1S-BE , one teacher will gain by misreporting his preferences. Assume that there are five teachers $t_1, \ldots, t_5$ and five schools $s_1, \ldots, s_5$.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

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

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

$$\mu_0 = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 & t_5 \\ s_1 & s_2 & s_3 & s_4 & s_5 \end{pmatrix}$$

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 node to another if the teacher in the origin of the arrow prefers the school in the pointed node. Theses arrows are not actual arrows of the graph associated with 1S-BE and so cannot be used to select a cycle. These arrows 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 $\underline{C}$.

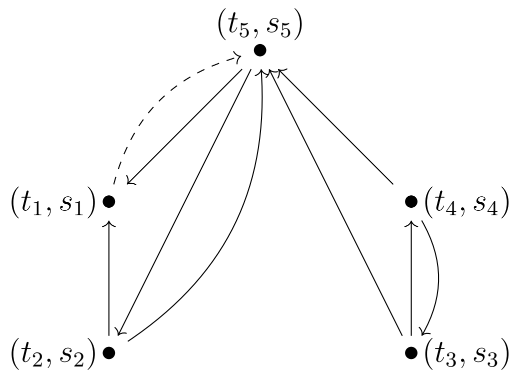
So we decompose the analysis for these two cases.

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

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:

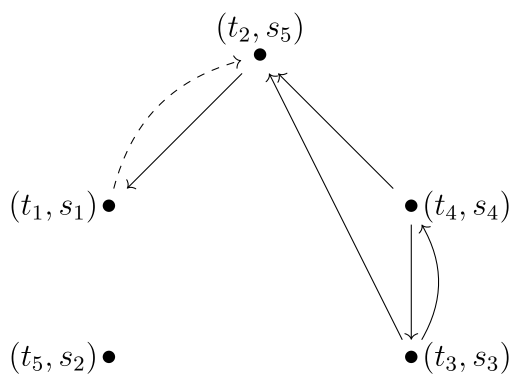
$$\bar{\mu} = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 & t_5 \\ s_1 & s_3 & s_4 & s_5 & s_2 \end{pmatrix}$$

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:



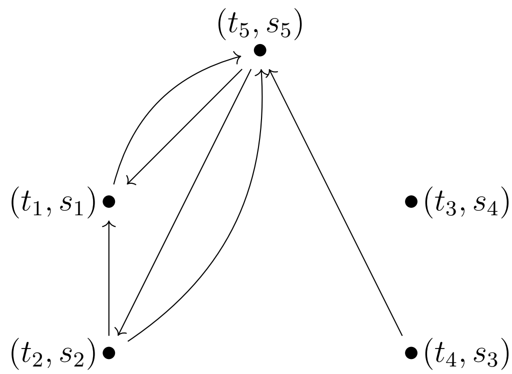
Now, there are two possible choices of cycles.

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, there is a unique cycle $(t_4, s_4) \leftrightarrows (t_3, s_3)$. Consider the new matching once this cycle is implemented. Teacher t_3 obtains his most favorite school. Hence, in the graph associated with 1S-BE starting from the new matching, node (t_1, s_1) will now point to node (t_2, s_5). In this graph, the only cycle is $(t_2, s_5) \leftrightarrows (t_1, s_1)$ and so t_2 is eventually matched 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:



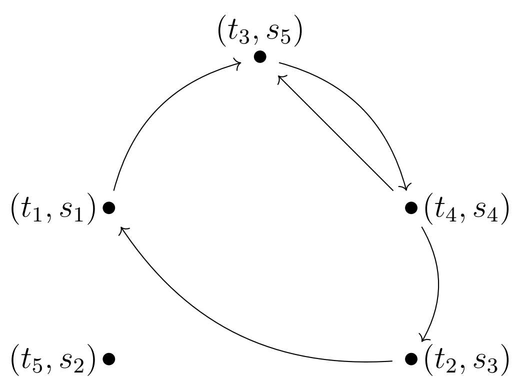
In this graph, there are three possible choices of cycles:

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 cycle $\overline{C}$ is selected under 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 selected:

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 when we selected cycle $\bar{C}$. So we can 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, each teacher but teacher t_4 gets his most preferred school. Hence, there is no more cycles in the new graph associated with 1S-BE. In particular, teacher t_4 gets matched to school s_3. Now, 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 arrows 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 we 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 select $\bar{C}$, 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 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 **bipartite graph** $G_b(n)$ consists of $2n$ vertices $V_1 \cup V_2$ (each of equal size) and edges $E \subset V_1 \times V_2$ across V_1 and V_2 (with no possible edges within vertices in each side). Random (bipartite) graphs can be seen as random variables over the space of (bipartite) graphs. We will see two asymptotic properties of random graphs: one based on the notion of perfect matchings, the other on that of independent sets.

A **perfect matching** of $G_b(n)$ is a subset E' of E such that each node in $V_1 \cup V_2$ is contained in a single edge of E'.

Lemma 10 (Erdös-Rényi) *Fix $p \in (0,1)$. Consider a "random graph" which selects a graph $G_b(n)$ with the following procedure. Each pair $(v_1, v_2) \in V_1 \times V_2$ is linked by an edge with probability p independently (of edges created for all other pairs). The probability that there is a perfect matching in a realization of this random graph tends to 1 as $n \to \infty$.*

The second important technical result is about so called independent sets. An **independent set** of $G(n)$ is $\bar{V} \subseteq V$ such that for any $(v_1, v_2) \in \bar{V} \times \bar{V}$, (v_1, v_2) is not in E.

Lemma 11 (Grimmett and McDiarmid (1975)) *Fix $p \in (0,1)$. Consider a "random graph" which selects a graph $G(n)$ with the following procedure. 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). Then,*

$$\Pr\left\{\exists \text{ an independent set } \bar{V} \text{ such that } |\bar{V}| \geq \frac{2\log n}{\log \frac{1}{1-p}}\right\} \to 0 \text{ as } n \to \infty.$$

E.2 Proof of Theorem 7

In the sequel, we fix μ_0 and 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 9 *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 9. Fix an arbitrary k and fix $\varepsilon > 0$. 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 $\xi_{t\mu_0(t')} > 1-\varepsilon$ and $\xi_{t'\mu_0(t)} > 1-\varepsilon$ and $\eta_{t'\mu_0(t)} > 1-\varepsilon$ and $\eta_{t\mu_0(t')} > 1-\varepsilon$. Then, in the random graph, each edge between $(t, \mu_0(t))$ and $(t', \mu_0(t'))$ is added independently with probability $\varepsilon^4 \in (0,1)$. Then, let $\hat{T}_k := \{t \in T_k \,\middle|\, \varphi(t) = \mu_0(t) \text{ and } U_t(\mu_0(t)) \leq u_k + 1 - \varepsilon \text{ and } V_{\mu_0(t)}(t) \leq 1 - \varepsilon\}$. It must be that $\{(t, \mu_0(t))\}_{t \in \hat{T}_k}$ is an independent set, or else if there is an edge $(t, \mu_0(t)), (t', \mu_0(t'))$ where $t, t' \in \hat{T}_k$ for some realization of the random graph, then

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

and similarly,

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

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 11 to get that $\Pr\left\{|\hat{T}_k| \geq \frac{2\log(|T_k|)}{\log\frac{1}{1-p}}\right\} \to 0$ as $n \to \infty$ and thus $\frac{|\hat{T}_k|}{|T_k|} \xrightarrow{p} 0$ as $n \to \infty$. Setting $\tilde{T}_k := \{t \in T_k \,|U_t(\mu_0(t)) \leq u_k + 1 - \varepsilon \text{ and } V_{\mu_0(t)}(t) \leq 1 - \varepsilon\}$, 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 $\varepsilon > 0$ 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 10 *There is a mechanism φ which is a selection of the BE algorithm such that for any k and any $\varepsilon > 0$ we have*

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

where $\bar{T}_k := \{t \in T_k \,|U_t(\varphi(t)) \geq u_k + 1 - \varepsilon\}$ and $\bar{S}_k := \{s \in S_k \,|V_s(\varphi(s)) \geq 1 - \varepsilon\}$.

Proof of Proposition 10. Fix $\varepsilon > 0$. We show that there exists a 2-IR mechanism ψ s.t. for each $k = 1, ..., K$, it matches each teacher $t \in T_k$ to a school in S_k and for each $\delta > 0$:

$$\Pr\left\{\frac{\left|\left\{t \in T_k \,|\xi_{t\psi(t)} \geq 1 - \varepsilon\right\}\right|}{|T_k|} > 1 - \delta\right\} \to 1$$

and

$$\Pr\left\{\frac{\left|\left\{s \in S_k \,|\eta_{\psi(s)s} \geq 1 - \varepsilon\right\}\right|}{|S_k|} > 1 - \delta\right\} \to 1$$

as $n \to \infty$ where we recall that $T_k := \mu_0(S_k)$. This turns out to be enough for our purpose. Indeed, consider the matching mechanism given by $\varphi := \text{BE} \circ \psi$ (i.e., the mechanism which runs BE on top of the assignment found by mechanism ψ). Since ψ is 2-IR so is φ. Hence, by construction, this must be a selection of BE which satisfies

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

as $n \to \infty$.

Fix $k = 1, ..., K$. Fix $\varepsilon_0 \in (0, \varepsilon)$. Further assume that ε_0 is small enough so that $(1 - \varepsilon_0)^2 > 1 - \delta$. Consider the set of pairs $(t, s) \in T_k \times S_k$ such that $s = \mu_0(t)$ and either t ranks s within his $\varepsilon_0|S_k|$ most favorite schools in S_k or s ranks t within his $\varepsilon_0|T_k|$ most favorite teachers in T_k. We eliminate these pairs from $T_k \times S_k$. Observing that the remaining set is a product set we denote it by $T_k^0 \times S_k^0$. Note that for each pair $(t, s) \in T_k \times S_k$ such that $s = \mu_0(t)$, there is a probability $(1 - \varepsilon_0)^2$ that both t ranks s outside his $\varepsilon_0|S_k|$ most favorite schools in S_k and s ranks t outside his $\varepsilon_0|T_k|$ most favorite teachers in T_k. Let us call E_{ts} this event. For each such (t, s) where $s = \mu_0(t)$ we denote $\mathbf{1}_{ts}$ for the indicator function which takes value 1 if the event E_{ts} is true and 0 otherwise. Hence, $|T_k^0| = \sum_{(t,s)\in T_k\times S_k : s=\mu_0(t)} \mathbf{1}_{ts}$. Thus, $|T_k^0|(= |S_k^0|)$ follows a Binomial distribution Bin($|T_k|$,$(1 - \varepsilon_0)^2$). By the law of large numbers, $\frac{|T_k^0|}{|T_k|} \xrightarrow{p} (1 - \varepsilon_0)^2$ which by assumption is strictly greater than $1 - \delta$. This proves that

$$\Pr\left\{\frac{|T_k^0|}{|T_k|} \geq 1 - \delta\right\} \to 1$$

and

$$\Pr\left\{\frac{|S_k^0|}{|S_k|} \geq 1 - \delta\right\} \to 1.$$

In the sequel, we condition w.r.t. a realization of the random set $T_k^0 \times S_k^0$ assuming that both $\frac{|T_k^0|}{|T_k|}$ and $\frac{|S_k^0|}{|S_k|}$ are greater than $1 - \delta$. Now, fix $\varepsilon_0' > 0$ and note that conditional on this, each teacher $t \in T_k^0$ draws randomly[44] in S_k^0 his $\varepsilon_0'|S_k^0|$ most favorite schools in S_k^0. Similarly, each school $s \in S_k^0$ draws randomly in T_k^0 its $\varepsilon_0'|T_k^0|$ most favorite teachers in T_k^0. We build a random bipartite graph on $T_k^0 \cup S_k^0$ where the edge $(t, s) \in T_k^0 \times S_k^0$ is added if and only if t ranks s within his $\varepsilon_0'|S_k^0|$ most favorite schools in S_k^0 and, similarly, s ranks t within its $\varepsilon_0'\,|T_k^0|$ most favorite teachers in T_k^0. This random bipartite graph can be seen as a mapping from the set of ordinal preferences into the set of bipartite graph $G_b(|T_k^0|)$. We denote this random graph by $\tilde{G}_b$. While Lemma 10 does not apply directly to this type of random graph, we will claim below that this random graph has a perfect matching, with probability approaching one as the market grows. Before stating and proving this result, we need the following lemma

Lemma 12 *With probability approaching one, for any teacher $t \in T_k^0$, any school $s \in S_k^0$ with which $\xi_{ts} \geq 1 - \frac{\varepsilon_0'}{2}$ must be within his $\varepsilon_0'|S_k^0|$ most favorite schools in S_k^0. Similarly, with*

[44]In the following, by randomly, we mean uniformly i.i.d.

probability approaching one, for any school $s \in S_k^0$, any teacher $t \in T_k^0$, with whom $\eta_{ts} \geq 1 - \frac{\varepsilon_0'}{2}$ must be within his $\varepsilon_0'|S_k^0|$ most favorite teachers in T_k^0.

PROOF. We prove the first part of the statement, the other part follows the same argument. Fix $t \in T_k^0$ and let E_t be the event that any school $s \in S_k^0$ with which $\xi_{ts} \geq 1 - \frac{\varepsilon_0'}{2}$ must be within his $\varepsilon_0'|S_k^0|$ most favorite schools in S_k^0. Let $X_t := \sum_{s \in S_k^0} \mathbf{1}_{\left\{\xi_{ts} \geq 1 - \frac{\varepsilon_0'}{2}\right\}}$ be the number of schools in S_k^0 with which teacher t enjoys an idiosyncratic payoff greater than $1 - \frac{\varepsilon_0'}{2}$. Observe that X_t follows a Binomial distribution $B(|S_k^0|, \frac{\varepsilon_0'}{2})$ (recall that ξ_{ts} follows a uniform distribution with support $[0, 1]$) and that $X_t \leq \varepsilon_0'|S_k^0|$ implies that E_t is true. Hence, we have to prove that $\Pr\{\exists t \in T_k^0 : X_t > \varepsilon_0'|S_k^0|\} \to 0$ as $n \to \infty$. In the sequel, we let Y_t be a Binomial distribution $B(|S_k^0|, 1 - \frac{\varepsilon_0'}{2})$, we have

$$
\begin{aligned}
\Pr\left\{\exists t \in T_k^0 : X_t > \varepsilon_0'|S_k^0|\right\} &\leq |T_k^0| \Pr\left\{X_t > \varepsilon_0'|S_k^0|\right\} \\
&= |T_k^0| \Pr\left\{Y_t \leq (1 - \varepsilon_0')|S_k^0|\right\} \\
&\leq |T_k^0| \exp\left\{-2|S_k^0|(\frac{\varepsilon_0'}{2})^2\right\} \to 0
\end{aligned}
$$

as $n \to \infty$, where the first inequality is by the union bound and the last one uses Hoeffding inequality. The limit result uses the fact that under our conditioning event, $|T_k^0| = |S_k^0| \geq (1 - \delta)|S_k| \to \infty$. ∎

We now move to our statement on the existence of a perfect matching in $\tilde{G}_b$.

Lemma 13 *With probability going to 1 as $n \to \infty$, the realization of $\tilde{G}_b$ has a perfect matching.*

PROOF. In our random environment, the state space, say Ω, can be considered as the set of all possible profiles of idiosyncratic shocks for teachers and schools, i.e., the space of all $\{\{\xi_{ts}\}_{ts}, \{\eta_{ts}\}_{ts}\}$. We denote by ω a typical element of that set. Let E be the event under which "For each $(t, s) \in T_k^0 \times S_k^0 : \xi_{ts} \geq 1 - \frac{\varepsilon_0'}{2}$ and $\eta_{ts} \geq 1 - \frac{\varepsilon_0'}{2}$ imply that both t ranks s within his $\varepsilon_0'|S_k^0|$ most favorite schools in S_k^0 and s ranks t within his $\varepsilon_0'|T_k^0|$ most favorite teachers in T_k^0". By Lemma 12, $\Pr(E) \to 1$. Now, let us build the following random graph on $T_k^0 \cup S_k^0$ where this time the edge $(t, s) \in T_k^0 \times S_k^0$ is added if and only if $\xi_{ts} \geq 1 - \frac{\varepsilon_0'}{2}$ et $\eta_{ts} \geq 1 - \frac{\varepsilon_0'}{2}$. Let us call this graph $\tilde{G}_b'$. So this time, $\tilde{G}_b'$ can be viewed as a mapping from the set of cardinal preferences into the set of bipartite graph $G_b(|T_k^0|)$. Let F be the event that the realization of $\tilde{G}_b'$ has a perfect matching. By Lemma 10, $\Pr(F) \to 1$. By definition, $E \cap F \subset \Omega$. Let us consider the set of all possible profiles of teachers and schools' ordinal preferences $\succ$ induced by states $E \cap F$ and let us denote this set by $\mathcal{P}$. Clearly, $\Pr(\mathcal{P}) \geq \Pr(E \cap F) \to 1$. Now, for each profile of preferences $\succ$ in $\mathcal{P}$, let $\tilde{G}_b(\succ)$ be the graph corresponding to $\tilde{G}_b$ when $\succ$ is the profile of realized preferences. We claim that for any $\succ$ in $\mathcal{P}$, $\tilde{G}_b(\succ)$ has a perfect matching. Indeed, let $\omega \in E \cap F$ be one state which induces $\succ$ (this is well defined by construction of $\mathcal{P}$). Because $\omega \in F$, the realization of $\tilde{G}'_b$ at profile ω has a perfect matching. In addition, because $\omega \in E$, the realization of $\tilde{G}'_b$ at profile ω is a subgraph of $\tilde{G}_b(\succ)$. We conclude that $\tilde{G}_b(\succ)$ has a perfect matching. Combining this result with the observation that $\Pr(\mathcal{P}) \to 1$, we get

$$\Pr\left\{\exists \text{ a perfect matching in } \tilde{G}_b\right\} \to 1$$

as $n \to \infty$, as claimed. ∎

Now, we build the mechanism ψ as follows. For each realization of ordinal preferences, (for each $k = 1, ..., K$) we build a graph on $T_k^0 \cup S_k^0$ as defined above, i.e., where the edge $(t, s) \in T_k^0 \times S_k^0$ is added if and only if t ranks s within his $\varepsilon'_0 |S_k^0|$ most favorite schools in S_k^0 and, similarly, s ranks t within its $\varepsilon'_0 |T_k^0|$ most favorite teachers in T_k^0. If there is a perfect matching, then under ψ, teachers in T_k^0 are matched according to this perfect matching while teachers in $T_k \backslash T_k^0$ remain at their initial assignment. If there is no perfect matching then under ψ, all teachers in T_k remain at their initial assignment. Assuming that $\varepsilon'_0 + \delta < \varepsilon_0$, we get that the mechanism built in that way is 2-IR.[45] To see this, consider a teacher t who does not get matched to his initial school. This means that t is matched to a school s given by a perfect matching of the random bipartite graph. By construction, this means that t ranks s within his $\varepsilon'_0 |S_k^0|$ most favorite schools in S_k^0. Hence, this means that s is within his $\varepsilon'_0 |S_k^0| + \delta |S_k|$ most favorite schools in S_k. Since $\varepsilon'_0 |S_k^0| + \delta |S_k| \leq (\varepsilon'_0 + \delta) |S_k| < \varepsilon_0 |S_k|$ and because $t \in T_k^0$ implies that $\mu_0(t)$ is not within t's $\varepsilon_0 |S_k|$ most favorite schools in S_k, we obtain that s is prefered by t to his initial assignment. Since a similar reasoning holds for schools, we obtain that ψ is 2-IR.

As we have shown, with probability approaching one, our bipartite graph actually has a perfect matching. Obviously, this perfect matching ensures that all teachers in T_k^0 and all schools in S_k^0 get matched to a partner within their $\varepsilon'_0 |S_k^0|$ favorite. This holds for any realization of the random set $T_k^0 \times S_k^0$ such that $\frac{|T_k^0|}{|T_k|}$ and $\frac{|S_k^0|}{|S_k|}$ are greater than $1 - \delta$. Thus, it holds conditional on the random sets $\frac{|T_k^0|}{|T_k|}$ and $\frac{|S_k^0|}{|S_k|}$ being greater than $1 - \delta$. Hence, this perfect matching ensures that all teachers in T_k^0 and all schools in S_k^0 get matched to a partner within their $(\varepsilon'_0 + \delta) |S_k|$ favorite in S_k and T_k respectively. Hence, under our conditioning event that

[45] This is without loss of generality because if $\Pr\left\{\frac{\left|\left\{t \in T_k | \xi_{t\psi(t)} \geq 1 - \varepsilon\right\}\right|}{|T_k|} > 1 - \delta\right\} \to 1$ then, $\Pr\left\{\frac{\left|\left\{t \in T_k | \xi_{t\psi(t)} \geq 1 - \varepsilon\right\}\right|}{|T_k|} > 1 - \delta'\right\} \to 1$ for any $\delta' > \delta$.

the random sets $\frac{|T_k^0|}{|T_k|}$ and $\frac{|S_k^0|}{|S_k|}$ are greater than $1-\delta$,

$$\Pr\left\{\frac{|\{t \in T_k\, |\psi(t) \text{ is within the } (\varepsilon_0' + \delta)\, |S_k| \text{ most favorite school in } S_k\}|}{|T_k|} > 1-\delta\right\} \to 1$$

and

$$\Pr\left\{\frac{|\{s \in S_k\, |\psi(s) \text{ is within the } (\varepsilon_0' + \delta)\, |S_k| \text{ most favorite teacher in } T_k\}|}{|S_k|} > 1-\delta\right\} \to 1$$

Given that the conditioning event has a probability approaching 1 as $n \to \infty$, this is even true without conditioning.

Now, without loss of generality, let us assume that δ is small enough so that $\varepsilon_0' + \delta < \varepsilon$. It remains to show that these $(\varepsilon_0' + \delta)\, |S_k|$ favorite partners in S_k (resp. T_k) yield an idiosyncratic payoff greater than $1-\varepsilon$. The following lemma completes the argument.

Lemma 14 *With probability going to* 1 *as* $n \to \infty$*, the* $(\varepsilon_0' + \delta)\, |S_k|$ *most favorite schools of each teacher in* T_k *yield an idiosyncratic payoff higher than* $1-\varepsilon$ *and the* $(\varepsilon_0' + \delta)\, |T_k|$ *most favorite teachers of each school in* S_k *yield an idiosyncratic payoff higher than* $1-\varepsilon$.

Proof. We show that with probability going to 1 as $n \to \infty$, the $(\varepsilon_0' + \delta)\, |S_k|$ most favorite schools of each teacher in T_k yield an idiosyncratic payoff higher than $1-\varepsilon$. The other part of the statement is proved in the same way. For each $t \in T_k$, let Z_t be the number of schools s in S_k for which $\xi_{ts} \geq 1-\varepsilon$. Note that if $Z_t > (\varepsilon_0' + \delta)\ |S_k|$ then t's $(\varepsilon_0' + \delta)\ |S_k|$ first schools in S_k must yield an idiosyncratic payoff higher than $1-\varepsilon$. Thus, it is enough to show that

$$\Pr\{\exists t \in T_k \text{ with } Z_t \leq (\varepsilon_0' + \delta)\ |S_k|\} \to 0$$

as $n \to \infty$. Observe that Z_t follows a Binomial distribution $B(|S_k|, \varepsilon)$ (recall that ξ_{ts} follows a uniform distribution with support $[0,1]$). Hence,

$$\begin{aligned}
\Pr\{\exists t \in T_k \text{ with } Z_t \leq (\varepsilon_0' + \delta)\ |S_k|\} \leq & \sum_{t \in T_k} \Pr\{Z_t \leq (\varepsilon_0' + \delta)\ |S_k|\} \\
= & \ |T_k| \Pr\{Z_t \leq (\varepsilon_0' + \delta)\ |S_k|\} \\
\leq & \ |T_k| \frac{1}{2} \exp\left(-2\frac{(|S_k|\varepsilon - (\varepsilon_0' + \delta)\,|S_k|)^2}{|S_k|}\right) \\
= & \ \frac{|T_k|}{2\exp\left(2\left(\varepsilon - (\varepsilon_0' + \delta)\right)^2 |S_k|\right)} \to 0
\end{aligned}$$

where the first inequality is by the union bound while the second equality is by Hoeffding's inequality. ∎

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

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

Proof of Proposition 11. 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 $T_k^0 := \{t \in T_k \left|\eta_{t\mu_0(t)} \leq 1 - \delta\right\}$.

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 $|T_k^0|$ 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 $\xi_{t\mu_0(t')} > 1 - \varepsilon$ and $\xi_{t'\mu_0(t)} > 1 - \varepsilon$ 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 $\varepsilon^2\delta^2 \in (0, 1)$. 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

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

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 11 to get that $\Pr\left\{|\bar{T}_k^0| \geq \frac{2\log(|T_k|)}{\log\frac{1}{1-p}}\right\} \to 0$ 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

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

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 coalition 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 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 Teacher-Optimal Block Exchange Algorithm

In the following lines, we define a class of mechanisms which are all selections of the BE algorithm and are strategy-proof. They all reduce to the TO-BE mechanism (as defined in the main text) in the one-to-one environment.

Given a matching μ and a set of school $S' \subseteq S$, we let $\text{Opp}(t, \mu, S') := \{s \in S' \,|t \succeq_s t' \text{ for some } t' \in \mu(s)\}$ 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 $\mu_0(t) \in \text{Opp}(t, \mu_0, S')$.

Now, fix an ordering over teachers $f : \{1, \ldots, |T|\} \to T$ which will be the index for our class of mechanisms.

- **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 $\text{Opp}(t, \mu(k-1), S(k-1)) = \text{Opp}(t, \mu_0, S(k-1))$, teacher t' has a lower priority than teacher t at school s' and teacher t' has the lowest ordering according to f among all teachers forming a pair with school s' and having a lower priority than t at s (i.e., $f(t') \leq f(t'')$ for all t'' such that $\mu(k-1)(t'') = s'$ and $t \succeq_s t''$). The directed graph so obtained is a directed graph with out-degree one 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 $k+1$.

Remark 4 *This class of mechanisms is still tightly connected to the top trading cycle mechanism. To see this, fix an ordering f and assume first that each teacher ranks schools outside his opportunity set below his initial assignment. Now, for each teacher define preferences over pairs $\{(t, \mu_0(t))\}_{t \in T}$ in the following way: for $s \neq s'$, (t, s) is strictly preferred to (t', s') if and only if s is strictly preferred to s'; in addition, (t, s) is strictly preferred to (t', s) if and only if $f(t) < f(t')$. We can consider this modified environment as a one-to-one environment where agents' preferences are strict. Top trading cycles is well-defined in this environment and coincides with the outcome of TO-BE (with the ordering f) defined in the previous paragraph.*

The 1-Sided Block Exchange Algorithm

In order to keep the property that the graph associated with 1S-BE is a supergraph of that of BE, we build on the previous generalization 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 toward t' at s'; or (2) t desires s' and t is ranked higher by s' than each teacher who both desires s' and does not block with s'.[46] 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 $k+1$.

G Empirical results

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.[47] Second, and perhaps more importantly, teachers can report "wide wishes"instead of reporting a precise school. A wide wish can be a geographic area such as a city, a group of cities, a department or an entire region. For instance, instead of ranking a school within city x, a teacher can report the city x in his ranking. These wide wishes make it possible for teachers to cover all schools within a region – if they wish to – which would not be possible otherwise due to the limit of 20 schools that applies. Then, when a city is ranked, the designer is free to assign the teacher to any school within that city. This is a way for teachers to signal their strong preference to be in a city x (irrespective of where they eventually end up in that city). While this may make sense from a design point of view, it makes ir more difficult for us to interpret the reported preferences.

25 067 teachers participate to the first phase of the assignment. We restrict the sample to the 49 subjects containing more than 10 teachers asking for a transfer : this restricts the sample to 20 808 teachers. We also remove from the sample all couples (1579 teachers) because of the specific treatment they receive in the assignment procedure.[48] Finally, only

[46] Note that here, teacher t may block with s' under condition (2). Thus, it is easy to see that if (1) is satisfied then (2) is satisfied as well. Hence, one could simplify the definition and suppress condition (1). We keep it just to have a parallel with the definition provided in the one-to-one environment.

[47] Teachers can rank up to 20 or 30 schools, depending on the region. In regions where they can rank a maximum of 20 schools, 10.79 % of the teachers rank 20 schools. In regions where teachers can rank up to 30 schools, the constraint is binding for less than 1 % of the teachers.

[48] Couples can jointly apply, in which case they have to submit two identical lists of regions to the central administration. A specific treatment is applied to the couple: one of the spouses will not be assigned a region if the other one does not get the same region. To achieve this, the central administration runs the algorithm once

teachers who have an initial assignment are kept in the sample. The final sample contains 10579 teachers.

and checks the region obtained by each spouse. If they don't obtain the same region, one must have obtained a region that is ranked lower in their common ranking (for instance rank 5). In that case, the ministry would delete all regions ranked higher than rank 5 from their common list of preferences and run the algorithm again on the modified list. This process is repeated until both spouses obtain the same regions. If this does not happen, they stay in their initial region.