Making Just School Assignments

Thayer Morrill[†]

December 2014

Abstract

It is well known that it is impossible for a mechanism to be strategyproof, Pareto efficient, and eliminate justified envy. However, little is known to what extent a strategyproof and efficient mechanism can limit justified envy. We define an assignment to be *unjust* if a student i is not assigned to a school a that she prefers to her own assignment, i has higher priority at a than one of the students assigned to a, and none of the students ranked higher at a than i are dependent on j. We prove that Top Trading Cycles is the unique mechanism that is strategyproof, efficient, and just. This demonstrates that any strictly stronger nothing of fairness is either unachievable by a strategyproof and efficient mechanism or else logically equivalent to justness in this class of mechanisms. We extend this characterization to the general case when schools may have arbitrary capacities by introducing the concept of reducibility.

[†]North Carolina State University. Email address: thayer_morrill@ncsu.edu. I am grateful to Umut Dur and Melinda Morrill for their helpful comments.

Key Words: Top Trading Cycles, School Choice, Assignment.
JEL Classification: C78, D61, D78, I20

1 Introduction

The school choice problem is one of the most important and well-studied problems in market design. The problem considers mechanisms for assigning students to public schools. Here, only students are considered to be strategic agents and the schools are treated as objects without preferences. However, a school board, which chooses the assignment algorithm, certainly has preferences over the resulting assignment. Ideally, the school board would like the student assignment to be both Pareto efficient and fair. Eliminating justified envy is widely used as the definition of a fair assignment,[1] but unfortunately, it is impossible for an assignment mechanism to be Pareto efficient and always eliminate justified envy.[2] Therefore a school board must choose between making an efficient assignment or an assignment that eliminates justified envy. Given this choice, the overwhelm-

[1]Student i is said to have *justified envy* if there is a school a such that i prefers a to her assignment, and i has higher priority at a than one of the students assigned to s. This is closely related to stability in the college admissions problem (Gale and Shapley, 1962). The critical difference is that in the school assignment problem, schools are treated as objects without preferences where as in the college admissions problem, schools are treated as agents with preferences over students. See the seminal paper Abdulkadiroglu and Sonmez (2003) for a detailed discussion of the similarities and differences between these two problems.

[2]See Example 1 for an assignment problem where the unique assignment eliminating justified envy is Pareto inefficient. This example was taken from Abdulkadiroglu and Sonmez (2003). The incompatibility of efficiency and stability was demonstrated in Roth (1982).

ing majority of school districts have chosen to eliminate justified envy.[3] Clearly making "fair" assignments is of central importance to school boards. Therefore, it is surprising that there has been relatively little discussion in the literature as to what constitutes a fair assignment.[4]

It is clear that if no student has justified envy, then an assignment is fair. What is less clear is whether or not every instance of justified envy means the assignment is unfair. The central question this paper addresses is to what extent is a strategyproof and efficient mechanism able to limit justified envy. Towards this aim, we introduce a weaker interpretation of fairness which balances respecting a student's priority at a school with respecting the assignments of students ranked higher than her at the school. Specifically, we consider an assignment unjust if a student i prefers a school a to her assignment, has higher priority at a than one of the students j that is assigned to a, and none of the the students ranked higher than i at a depend on j for her assignment. The key consideration is what it means for student i to depend on student j. We define i to *depend* on j if j, by submitting an alternative preference profile, is able to receive i's assignment. When i depends

[3]Cities that have adopted a version of the student-proposing Deferred Acceptance algorithm include New York City (Abdulkadiroglu et al. 2005b, 2009), Boston (Abdulkadiroglu et al. 2005a), and Chicago (Pathak and Sonmez, forthcoming). Denver began using DA in 2012. Recently, DA has been adopted by all local authorities in England (Pathak and Sonmez, forthcoming). The only school district we know of that has implemented Top Trading Cycles is New Orleans.

[4]This paper considers fairness from an ex-post perspective. A number of papers have considered fairness in assignment mechanisms from an ex-ante perspective. In that context, a mechanism is typically interpreted as fair if it is symmetric: two agents who submit the same preference profile receive the same distribution over objects. An alternative fairness notion in this environment is envy-freeness: each agent first-order stochastically prefers her allocation to the allocation of any other agent. See Bogomolnaia and Moulin (2001), Che and Kojima (2010), and Liu and Pycia (2013) among others for a more detailed discussion of fairness in this environment.

on j, then changing j's assignment has the potential to change i's assignment. Therefore, if i has justified envy of j at a, but a higher ranked student k depends on j, then we do not honor i's objection as it has the potential to harm a higher ranked student. Intuitively, we take the stance that an objection should be honored only if doing so has no potential to harm any of the higher ranked students.

Our main result is to demonstrate that there is no strictly stronger fairness condition than justness that is compatible with strategyproofness and efficiency. We do this by proving that Top Trading Cycles (hereafter TTC) is the unique mechanism that is strategyproof, efficient, and just. This demonstrates that any strictly stronger fairness concept is either impossible to achieve by a strategyproof and efficient mechanism, or else it is equivalent to justness in this class of mechanisms.

This characterization of TTC also provides a new way of understanding an important assignment mechanism. TTC does not eliminate justified envy, but there is no strategyproof and efficient mechanism that is fairer than TTC. Note that we do not claim that TTC is the fairest possible mechanism or that justness is the strongest possible criterion compatible with strategyproofness and efficiency. There may be conditions logically independent from justness that are satisfied by alternate mechanisms. This characterization holds when schools have a capacity of one. We extend this characterization to the general case when schools may have arbitrary capacities by introducing the concept of reducibility. Intuitively, a problem is reducible if large problems can always be separated into smaller subproblems. We show that Top Trading Cycles is the only mechanism that is strategyproof, efficient, just, and reducible.

1.1 Relation to the Literature

Our paper contributes to a substantial literature studying the extent to which fairness, efficiency, and strategyproofness are compatible for assignment problems. In addition to introducing the deferred acceptance algorithm, Gale and Shapley (1962) famously prove that among stable assignments for the college admissions problem, the assignment made by the student-proposing deferred acceptance algorithm Pareto dominates all other stable assignments from the perspective of the students. This result was applied to the school assignment problem first by Balinksi and Sonmez (1999) and then by Abdulkadiroglu and Sonmez (2003). Since only the students have preferences in the school assignment problem, these papers point out that there exists a Pareto optimal fair assignment. However, Abdulkadiroglu and Sonmez (2003) demonstrate that there is a fundamental tension between efficiency and fairness in the school assignment problem. It is impossible for a mechanism to be both fair and efficient. Kesten (2010) demonstrates that this tension between efficiency and fairness is exasperated when we restrict ourselves to strategyproof mechanisms. Specifically, he demonstrates that there is no strategyproof mechanism that always selects a fair and efficient assignment even when one exists. Abdulkadiroglu, Pathak, and Roth (2009) analyze both theoretically and empirically the efficiency loss associated with strategyproofness and fairness. In particular, they demonstrate that there exists no strategyproof mechanism that Pareto improves on the student proposing Deferred Acceptance algorithm with single tie breaking. In their analysis, 1.9% of the students could be matched to a school they strictly prefer without harming others in an alternative fair assignment. A further 5.5% of the students could be matched to a school they strictly prefer without harming others if we do not impose fairness.

The above papers consider deterministic mechanisms, but the interplay between fairness, efficiency, and strategyproofness has also been studied in the random environment. Here, symmetry is widely used as the definition of a fair mechanism. Bogomolnaia and Moulin (2001) demonstrate that no symmetric, strategyproof mechanism can be efficient. For example, a uniform randomization over TTC is strategyproof, symmetric, and ex-post efficient, but it is not ex-ante efficient. However, Kojima and Manea (2010b) and Che and Kojima (2010) demonstrate that this impossibility does not hold asymptotically in large markets. Strikingly, Liu and Pycia (2013) demonstrate that all asymptotically efficient, symmetric, and asymptotically strategyproof mechanisms converge to the same allocation. Of interest to the current paper, they demonstrate that uniform randomizations over TTC are asymptotically strategyproof, efficient, and symmetric (fair).

There have also been several recent papers that characterize TTC. Most closely related to this paper is Morrill (2013b) which demonstrates that when objects have unit capacities, TTC is the unique mechanism that is strategyproof, efficient, satisfies mutual best, and independent of irrelevant rankings.[5] The current paper has several advantages over Morrill (2013b). First, by characterizing TTC only in terms of strategyproofness, efficiency, and a fairness condition, our characterization provides a simpler answer as to what differentiates TTC from all other strategyproof and efficient mechanisms. Second, mutual best is a very weak fairness condition. While it is interesting from a characterization standpoint how weak of a condition is required to characterize the assignment, a policy maker is much more interested in how strong a notion of fairness that TTC satisfies. Finally, and most importantly, our characterization extends to the general case where objects may be assigned to multiple agents. The characterization in Morrill (2013b) does not

[5]Intuitively, a mechanism satisfies independence of irrelevant rankings if, when a student's ranking at a school never affects its own assignment, then it does not affect other students' assignments either. For a more precise definition, see Morrill (2013b).

as TTC does not satisfy independence of irrelevant rankings when objects have capacities greater than one.

Also closely related to the current paper are characterizations provided in Abdulkadiroglu and Che (2010). This was the first paper to study what distinguishes TTC from other Pareto efficient and strategyproof assignment mechanisms. They demonstrate that TTC is the only such mechanism that recursively respects top priorities. A just assignment respects top priorities, but justness is not a recursive concept and therefore is quite different from recursively respecting top priorities. Dur (2013) provides the first characterization of the general case. He shows that TTC is the unique mechanism satisfying Pareto efficiency, strategyproofness, mutual best, weak consistency, and resource monotonicity for top-ranked students.[6]

2 Model

We consider a finite set of agents $I = \{1, \ldots, n\}$ and a finite set of objects $O = \{a, b, c, \ldots\}$. Each agent $i \in I$ has a complete, irreflexive, and transitive preference relation P_i over $O \cup \{\emptyset\}$. $\emptyset$ represents an agent being unassigned, and there is no limit to the number of agents that may be assigned to $\emptyset$. aP_ib indicates

[6]More generally, this paper contributes to the growing literature on characterizations of assignment mechanisms. Papai (2000) characterizes hierarchical exchange rules, a general class of exchange rules which includes TTC. Pycia and Unver (2014) characterize a further generalization of hierarchical exchange rules called trading cycles. Kojima and Manea (2010a) characterize DA for some priorities of the objects. Morrill (2013a) extends this characterization to all substitutable priorities. Kojima and Unver (2010) characterize the Boston assignment mechanism. For the housing allocation problem with existing tenants, Sonmez and Unver (2010) provide a characterization of the you request my house-I get your turn mechanism introduced by Abdulkadiroglu and Sonmez (1999).

that i strictly prefers object a to b. Given P_i, we define the symmetric extension R_i by aR_ib if and only if aP_ib or $a = b$.

The capacity of each object $a \in O$ is given by q_a, and we set $q = \{q_a | a \in O\}$. Each school a has a complete, irreflexive, and transitive priority rule $\succ_a$ over I. In particular, $i \succ_a j$ is interpreted as agent i having a higher priority for object a than agent j. We define $\succeq$ analogously to our definition of R.

We let $P = (P_i)_{i \in I}$, $\succ = (\succ_a)_{a \in O}$, $P_{-I'} = (P_i)_{i \in I \setminus I'}$, and $\succ_{-O'} = (\succ_a)_{a \in O \setminus O'}$. Throughout, I, O, and the quotas q are fixed, and we define the assignment problem by $(P, \succ)$.

An **allocation** is a function $\mu : I \to O \cup \{\emptyset\}$ such that for each $a \in O$, $|\{i \in I | \mu(i) = a\}| \leq q_a$ where q_a is the capacity of a. In a slight abuse of notation, for a set of agents $I' \subset I$, we define $\mu(I') = \{a \in O | \exists i \in I' \text{ such that } \mu(i) = a\}$, and set $\mu(a) = \{i \in I | \mu(i) = a\}$. Given allocations μ and μ', we say $\mu R \mu'$ if $\mu(i) R_i \mu'(i)$ for every $i \in I$.

An allocation is **Pareto efficient** if there does not exist another allocation ν such that $\nu(i) R_i \mu(i)$ for every $i \in I$ and $\nu(i) P_i \mu(i)$ for some i.

We denote by $\mathcal{R}$, $\mathcal{C}$, and $\mathcal{A}$ the sets of all possible preference relationships, priority rules, and allocations, respectively. An **allocation mechanism** is a function $\phi : \mathcal{R} \times \mathcal{C} \to \mathcal{A}$. A mechanism ϕ is **strategyproof** if reporting true preferences is each agent's dominant strategy. That is:

$$\phi(R, \succ)(i) R_i \phi(R'_i, R_{-i}, \succ)(i)$$

for all R, $\succ$, $i \in I$, and R'_i. For notational convenience, we will typically fix the priority rule $\succ$ and denote the mechanism $\phi(R, \succ)$ by $\phi(R)$.

Abdulkadiroglu and Sonmez (2003) give detailed descriptions of TTC and DA. Given strict preferences of students and strict priority lists for schools, TTC assigns students to schools according to the following algorithm. In each round, each student points to her most preferred remaining school, and each school with available capacity points to the remaining student with highest priority. As there are a finite number of students, there must exist a cycle $\{o_1, i_1, \ldots, o_K, i_K\}$ such that each o_j and i_j points to i_j and o_{j+1}, respectively (with $o_{K+1} \equiv o_1$). For each cycle, student i_j is assigned to object o_{j+1}, i_j is removed, and the capacity of o_{j+1} is reduced by one. When a school has no remaining capacity, it is removed. The algorithm terminates when all students are assigned or no school has any available capacity. For any $R \in \mathcal{R}$, $\succ \in \mathcal{C}$, the mechanism $TTC(R, \succ)$ outputs the allocation made by TTC.

The student proposing version of DA is defined as follows. In the first round, each student proposes to her most preferred school. Each school tentatively accepts students up to its capacity and rejects the lowest priority students beyond its capacity. In every subsequent round, each student rejected in the previous round proposes to her most preferred school that has not already rejected her. Each school tentatively accepts the highest priority students up to its capacity and rejects all others. The algorithm terminates when there are no new proposals and tentative assignments are made final. Roth and Sotomayor (1992) is an excellent resource for the properties of DA.

3 Just Assignments

Eliminating justified envy is the notion of fairness that is typically considered by the literature. A student i has **justified envy** in assignment μ if there is a school a and a student j such that $aP_i\mu(i)$ and $i \succ_a j$ where $\mu(j) = a$. An assignment **eliminates justified envy** if no student has justified envy. It is well known that the assignment made by DA not only eliminates justified envy but Pareto dominates all other assignments that eliminate justified envy. However, Example 1 demonstrates that DA is not efficient. It is well known that TTC always makes an efficient assignment. However, Example 1 also demonstrates that TTC does not eliminate justified envy. Therefore, DA is typically interpreted as the fair assignment mechanism, and TTC it typically interpreted as the efficient mechanism.

Consider the following classic example from Roth (1982) and applied to the school assignment problem by Abdulkadiroglu and Sonmez (2003).

Example 1. There are three students i, j, k, and three schools a, b, c, each of which has a capacity of one. Consider the following preferences and priorities where P denotes the preferences of students and $\succ$ the priorities of schools.

P_i	P_j	P_k	$\succ_a$	$\succ_b$	$\succ_c$
b	a	a	i	j	j
a	b	b	k	i	i
c	c	c	j	k	k

There is only one assignment that eliminates justified envy:

$$\begin{pmatrix} i & j & k \\ a & b & c \end{pmatrix}$$

We label this assignment μ. However, μ is Pareto-dominated by the following assignment which we label λ:

$$\begin{pmatrix} i & j & k \\ b & a & c \end{pmatrix}$$

Under assignment λ, k has justified envy of j: aP_kc and $k \succ_a j$.

It seems clear that the highest ranked student at school s should either be assigned to s or a school she prefers to s.[7] If we agree to never violate these priorities, then in Example 1, k's claim to a is substantially weakened. Since i has the highest priority at a and the only school she prefers to a is b, i must be assigned to a or b. Similarly, j must be assigned to a or b. Therefore, k is never assigned to a so long as we honor the top priorities. The only way we can honor k's objection and to honor top priorities, is to assign i to a. Note that k's objection is based on her priority at a, but this results in an agent with even higher priority at a being harmed.

This motivates a new fairness concept we call justness. Intuitively, when a student raises an objection based on justified envy, we will allow the students with even higher priority at that object to veto the objection. In particular, suppose i has justified envy of j at object a. Before enforcing this objection, we first check whether changing j's assignment may possibly harm any of the students ranked even higher than i at a. If so, we "veto" the objection.[8]

Definition 1. Given an assignment mechanism ϕ, an agent i **depends** on agent j at preference profile R if there exist a R'_j such that $\phi(R'_j, R_{-j})(j) = \phi(R)(i)$.

[7] Abdulkadiroglu and Che (2010) introduce this concept and call it *respecting top priorities*.

[8] This condition is most natural when we consider strategyproof mechanisms and when objects have a capacity for one student.

Suppose j is assigned to school a and some student has justified envy of j. Let i be the highest ranked student with justified envy of j. Note that all of the students ranked higher than i at a strictly prefer their assignment to a. If none of these students depend on j, then they should have no objection to i being assigned to a. In this case, we consider it unjust if j is assigned to a instead of i. However, if one of these students depends on j, then it is not clear that we can change j's assignment without harming this student. In this case, we take the conservative position of not allowing such an objection.

Definition 2. Given a preference profile R, the assignment $\mu = \phi(R)$ is **unjust** if there exists a student i and school a such that

1. $aP_i\mu(i)$
2. $i \succ_a j$ where $\mu(j) = a$
3. For all students k such that $k \succ_a i$, k does not depend on j.

ϕ is **just** if it never makes an unjust assignment.

Since an instance of justified envy is necessary for an assignment to be unjust, clearly if an assignment has no justifiable envy, then it is just. Therefore, eliminating justified envy is a strictly stronger fairness concept than justness. In particular, since DA eliminates justified envy, it is just.

For intuition, consider whether or not TTC makes a just assignment in Example 1. Student k has justified envy of j at school a. However, if j ranks b first, then TTC will assign j to b. Therefore, i depends on j. Therefore, this instance of justified envy does not violate justness, and indeed, TTC makes a just assignment.

An assignment respects top priorities if the highest ranked student at each school s weakly prefers her assignment to s. Note that if an agent has the highest priority, then trivially there cannot be an agent with higher priority and so our additional restriction has no bite. Therefore, a just assignment respects top priorities.

The Boston mechanism is a common mechanism observed in practice.[9] In this mechanism, in the first round, each school only considers the students that have listed it as their first choice. Among these students, a school accepts the student with the highest priority and rejects all others.[10] In round k, each remaining student applies to the k^{th} school on her list. Each school with available capacity accepts the highest ranked student that applies. All other students are rejected. The algorithm terminates when all students have been assigned. Consider the case where i has highest priority at school a, j has highest priority at b, and k has highest priority at c. Suppose i ranks b first, a second and c third; j ranks b first; and k ranks a first. In this case, i is rejected by b in the first round. Since a accepts k in the first round, i is rejected by a in the second round. And ultimately i is assigned to c in the third round. Therefore, the Boston mechanism does not respect top priorities as i has highest priority at a yet is assigned to a school she finds inferior to a. Therefore, the Boston Mechanism is not just.

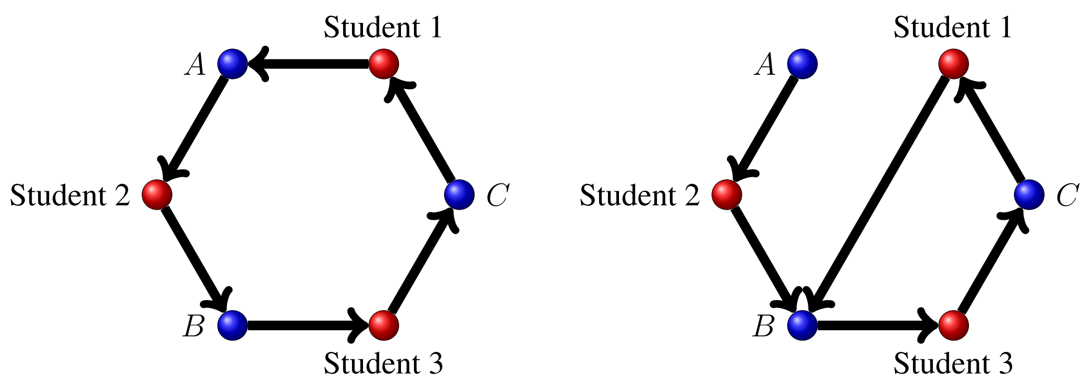
A significant practical objection to eliminating justified envy is that it is impossible for a mechanism to be Pareto efficient and eliminate justified envy. The next Lemma demonstrate that this is not the case with justness. TTC is strategyproof, Pareto efficient, and just.

Lemma 1. Top Trading Cycles is just.

[9] See Abdulkadiroglu and Sonmez (2003) for a detailed discussion.

[10] In general, if a school has capacity q, then it accepts up to the q highest ranked students.

Figure 1: Each agent in a trading cycle is dependent on the other agents in the cycle. For example, here Student 1 will receive Student 2's assignment if she points at it.



Proof. Figure 1 gives the intuition. In any trading cycle, all of the students in the cycle depend on each of the other students in the cycle. So if student i has justified envy of student j, then j was assigned in an earlier round than i. In that cycle, the object i envies is pointing to a student who is higher ranked than i and who depends on j. Therefore, this is not a violation of justness. More formally, consider any priorities and capacities of objects. Let R be any preference profile of the students. Suppose for some student i and school a that $a \; P_i \; TTC(R)(i)$ and that $i \succ_a j$ where $TTC(R)(j) = a$. Let $\{a, j_1, a_2, j_2, \ldots, a_n, j_n = j\}$ be the cycle in which TTC assigned a to j. Since i prefers a to her assignment, i is not assigned until after a has been assigned to capacity. In particular, $j_1 \succ_a i$. However, if j changed its preferences to R'_j where she ranks a_2 first, then she will be assigned a_2. Therefore, j_1 depends on j. $\square$

The following example is included both to provide intuition on justness and to demonstrate that strategyproofness is independent from efficiency and justness.

Example 2. Consider the following variation of TTC. For simplicity, we assume that the number of students equals the number of schools although the algorithm can be easily generalized. We will allow a school to deem a student unacceptable. Initially all students are acceptable at all schools. Each student points to her favorite school. Each school points to the highest-ranked acceptable student. Suppose a student i is the highest ranked student at more than one school. Let a be i's favorite school among those at which she is ranked first. We keep i's priority at a the same; however, for any other school b at which i has the highest priority, b now deems i unacceptable and points to what student is now its highest-ranked acceptable student. We iterate this process until no student has more than one school pointing at her.[11] As in TTC, there must exist at least one cycle. We process all cycles and then repeat until all students have been assigned. Note that i will not have justified envy at any school b that has declared i unacceptable. For b to declare i unacceptable, a school a that i prefers to b must have been pointing at i. Therefore, i is assigned to a or a school i prefers to a; therefore, she does not envy any student assigned to b. This algorithm is efficient since in each step the students who are assigned receive their favorite school with available capacity. This mechanism is also just for the same reason that TTC is just. The students in a cycle all depend on each other, and therefore if i has justified envy of j at object a, then a student ranked higher than i at a is dependent on j (in particular, the student that a is pointing to in the cycle involving j). Note that although this algorithm is efficient and just, it is not strategyproof. For example, consider the following

[11] Since we have assumed the number of schools equals the number of students, it is never the case that a school finds all remaining students unacceptable. However, if there are more schools then students (or in the general case when the total school capacity is greater than the number of students), such a school would be removed at that point.

agents, objects, and preferences. There are three students i, j, k, and three schools a, b, c, each of which has a capacity of one. The priorities and preferences are as follows:

R_i	R_j	R_k	$\succ_a$	$\succ_b$	$\succ_c$
b	a	a	i	j	j
a	c	b	k	k	i
c	b	c	j	i	k

Initially both b and c rank j first. Since j prefers, c to b, c points to j and we change b's priorities so that j is unacceptable. Now a points to i, b points to k, and c points to j. There is one cycle: $\{i, b, k, a\}$. Therefore, j is assigned to c. However, if j submits preferences $R'_j : a, b, c$, then j will be assigned to a.

A natural question to ask is whether or not it is possible to allow more objections than justness and still be strategyproof and Pareto efficient. We first demonstrate that when the capacity of each object is at most one, then it is not necessary to search for a stronger condition. TTC is the only strategyproof and efficient mechanism that is just, so any stronger condition could only be satisfied by TTC.

Theorem 1. Suppose each object's capacity equals one. Then the only mechanism that is strategyproof, efficient, and just is Top Trading Cycles.

Proof. It is well known that TTC is strategyproof and efficient. Lemma 1 establishes that TTC is just. Suppose for contradiction that there is a mechanism ϕ that is strategyproof, efficient, but there exists a $R \in \mathcal{R}$, $\succ \in \mathcal{C}$, and $i \in I$ such that $\phi(R, \succ)(i) \neq TTC(R, \succ)(i)$. We fix $\succ$ for the remainder of the proof.

For clarity, we consider an implementation of TTC where we process cycles one at a time. TTC is independent of the order in which cycles are processed. For each R such that $TTC(R) \neq \phi(R)$, we choose the order of cycles so that we assign an agent differently under TTC and ϕ as soon as possible. Specifically, define $f(R)$ to be the minimum number of cycles that must be processed before we assign an agent who receives a different assignment under TTC and ϕ. Among the preference profiles where TTC and ϕ are not equal, choose R such that $f(R)$ is minimal.

Let $C = \{o_1, i_1, o_2, i_2, \ldots, o_K, i_K\}$ be the earliest possible cycle where an agent is assigned differently under TTC and ϕ, and without loss of generality, assume $\phi(R)(i_K) \neq TTC(R)(i_K)$. Let I^* be the agents assigned in earlier cycles of TTC where $I^* = \emptyset$ if C is the first cycle. For $i_j \in C$ define $R_j^1 := o_{j+1}, o_j, \emptyset$ and define $R_j^2 := o_j, \emptyset$.

Claim 1: For any $i_j \in C$, $\phi(R_j^2, R_{I^*}, R'_{-j})(i_j) = o_j$ where R'_{-j} are any preferences of the agents $I \setminus (I^* \cup \{j\})$.

Suppose for contradiction that $\phi(R_j^2, R_{I^*}, R'_{-j})(i_j) \neq o_j$ for some $i_j \in C$. Since i_j is not assigned to o_j, i_j must be unassigned as otherwise leaving i_j unassigned would be a Pareto improvement. Moreover, if no student is assigned to o_j, then changing i_j's assignment to o_j would be a Pareto improvement. Therefore, there must be some student k such that $\phi(R_j^2, R_{I^*}, R'_{-j})(k) = o_j$. First, we show that $k \notin I^*$. Since $TTC(R_j^2, R_{I^*}, R'_{-j})(i) \neq o_j$ for any $i \in I^*$ and each $i \in I^*$ is processed in an earlier cycle than C, if $k \in I^*$, then k would be an agent assigned differently under TTC and ϕ in an earlier cycle than C. This would contradict the minimality of $f(R)$. Therefore, i_j has justified envy of k since i_j has the highest priority at o_j of any student not in I^* and therefore higher priority at o_j than k. Second, we show that no $i \in I^*$ is dependent on k. In fact, we show something stronger: for every $i \in I^*$ and every $l \in I$, if i is dependent on l, then $l \in I^*$. In words, no student in I^* is dependent on a student not in I^*. Consider any $i \in I^*$. i's assignment in TTC does not depend on the report of any student assigned in a later cycle. Therefore, if one of these students changed her report and it changed i's assignment under ϕ, then TTC and ϕ would assign i to different schools. Since each $i \in I^*$ is assigned in an earlier cycle than C, this would again contradict the minimality of $f(R)$. Therefore, i is not dependent on this student. In particular, no student in I^* is dependent on k. Since i_j has justified envy of k and no agent ranked higher at o_j than i_j is dependent on k, assigning k to o_j is unjust, a contradiction.

Claim 1 implies that for any $i_j \in C$, any R'_{i_j}, and any R'_{-j}, $\phi(R'_{i_j}, R_{I^*}, R'_{-j})(i_j)\ R'_{i_j} o_j$. Otherwise, i_j could strictly improve her assignment by reporting R^2_j which would violate strategyproofness. Therefore, for each $i_j \in C$ and any preferences R'_{-j},

$$\phi(R^1_{i_j}, R_{I^*}, R'_{-j})(i_j) \in \{o_{j+1}, o_j\}. \tag{1}$$

Since $\phi(R)(i_K) \neq o_1$, it cannot be that $\phi(R^1_{i_K}, R_{-i_K})(i_K) = o_1$ or else i_K could profitably misreport her preferences. Therefore, by Eq. (1), $\phi(R^1_{i_K}, R_{-i_K})(i_K) = o_K$. As o_K may only be assigned once, $\phi(R^1_{i_K}, R_{-i_K})(i_{K-1}) \neq o_K$. Since $\phi(R^1_{i_K}, R_{-i_K})(i_{K-1}) \neq o_K$, strategyproofness implies $\phi(R^1_{i_{K-1}}, R^1_{i_K}, R_{-\{i_{K-1}, i_K\}})(i_{K-1}) \neq o_K$ or else i_{K-1} could profitably misreport her preferences when her true preferences are $R_{i_{K-1}}$ and the other agents report preferences $(R^1_{i_K}, R_{-\{i_{K-1}, i_K\}})$. Therefore, by Eq. (1) $\phi(R^1_{i_{K-1}}, R^1_{i_K}, R_{-\{i_{K-1}, i_K\}})(i_{K-1}) = o_{K-1}$. Recursively applying this logic, we find that $\phi(R^1_C, R_{-C})(i_1) = o_1$. Eq. (1) and the fact that each object may be assigned only once imply that for every $i_j \in C$,

$$\phi(R^1_C, R_{-C})(i_j) = o_j \tag{2}$$

However, Eq (2) leads to a contradiction. Since $\phi(R^1_C, R_{-C})(i_j) = o_j$ for every

$i_j \in C$, then ϕ is inefficient as we can reassign each $i_j \in C$ to o_{j+1}, leave all other assignments unchanged, and Pareto improve $\phi(R^1)$. Therefore, $\phi(R, \succ)(i) = TTC(R, \succ)(i)$ for every $i \in I_1(R, \succ)$.

$\square$

DA is strategyproof and just but not efficient. Example 2 provides an algorithm that is efficient and just but is not strategyproof. A serial dictatorship is strategyproof and efficient but is not just. Therefore, the criteria in Theorem 1 are independent.

When schools have a capacity greater than one, then TTC is no longer the unique strategyproof, efficient, and just mechanism. For example, consider Clinch and Trade which was introduced in Morrill (2014b). Clinch and Trade is a variation on TTC. Each round consists of two parts. In the clinching phase, if the student has one of the q_a highest priorities at her most preferred object a, then we assign the agent and remove her. We iterate the clinching process until no student is able to clinch her assignment.[12] Next, in the pointing phase each agent points to her most preferred object with available capacity. As in TTC, we assign all cycles, remove agents, and adjust the capacities of objects accordingly. The algorithm terminates when all students have been assigned or no school has available capacity. The next lemma demonstrates that Clinch and Trade is also just.

Lemma 2. Clinch and Trade is just.

Proof. Consider any priorities and capacities of objects. Let R be any preference

[12]In the first round, all students are involved in the clinching phase. However, a student only participates in the clinching phase of round $k > 1$ if the school she was pointing to in round $k - 1$ no longer has available capacity. This restriction is necessary to preserve strategyproofness.

profile of the students. Let μ to be the assignment made by Clinch and Trade. Suppose $aP_i\mu(i)$ for some student i and school a, and let j be a student such that $\mu(j) = a$. If j was assigned via clinching, then j must have higher priority at a then i. Otherwise, let $\{a, j_1, a_2, j_2, \ldots, a_n, j_n = j\}$ be the cycle in which Clinch and Trade assigned a to j. Since i prefers a to her assignment, i is not assigned until after a has been assigned to capacity. In particular, $j_1 \succ_a i$. However, if j changed its preferences to R'_j where she switches the ordering of a and a_2, then she will be assigned a_2. Therefore, j_1 is dependent on j. Therefore, this instance of justified envy does not violate justness. $\square$

An important point to note is that for general capacities, we can no longer consider TTC to be “maximally” fair. It is an open question which assignment procedure has a minimal number of instances of justified envy when the object capacities are greater than one. However, Morrill (2014a) uses simulations to demonstrate that Clinch and Trade and a TTC variant called Prioritized Trading Cycles perform better than TTC on average.

TTC is an iterative algorithm. We assign the seats at schools one at a time, and whenever the capacity of a school is greater than one, we are always able to assign a set of students and reduce the size of the problem. This is a rather appealing feature of an algorithm. When a problem is complex, we are able to identify a group whose assignments have already been determined, assign them, and consider the remaining students separately. We are therefore always able to reduce a large problem to a simpler problem. This feature is enough to characterize TTC for arbitrary capacities. It is the only strategyproof, Pareto efficient, and just algorithm that is reducible in this way.

For notational convenience, we fix the set of objects O and the priorities of the objects $\succ$ over the agents I. It is understood that when we consider a subgroup of students $J \subset I$, that object a's priorities are the induced priorities $\succ_a^J$ where $j_1 \succ_a^J j_2$ if and only if $j_1 \succ_a j_2$ for any $j_1, j_2 \in J$. For the remainder of the paper, we denote an assignment problem $(I, O, R, \succ, q)$ by (I, R, q). Given two sets of students $J, K \subset I$ such that $J \cap K = \emptyset$, and given two assignments $\mu : J \to O$ and $\mu' : K \to O$, we define the assignment $\lambda = \mu \wedge \mu' : J \cup K \to O$ by $\lambda(j) = \mu(j)$ for each $j \in J$ and $\lambda(k) = \mu'(k)$ for each $k \in K$.

We want to be able to identify a group of students that we can assign and remove from consideration in order that we can simplify the problem. A key point is that the first group's assignments must not depend on the preferences of the remaining students.

Definition 3. Given a mechanism ϕ, (I^*, R_{I^*}, q^*), is a **dominant subproblem** for an assignment problem (I, R, q) if

$$\phi(I, R_{I^*}, R'_{I \setminus I^*}, q) = \phi(I^*, R_{I^*}, q^*) \wedge \phi(I \setminus I^*, R'_{I \setminus I^*}, q - q*)$$

where $0 < q^* < q$ and $R'_{I \setminus I^*}$ are any preferences of the students $I \setminus I^*$. We call I^* a **dominant subgroup**.

Dominant subproblems are not unique to TTC. For example, in DA an agent with highest priority at her favorite object forms a dominant subgroup. Such an agent's assignment is independent of the preferences of any other agent. Similarly, in a serial dictatorship, the j highest dictators form a dominant subgroup. A key distinguishing feature of TTC is that we are always able to find a dominant subgroup whenever the assignment problem is nontrivial.[13] We demonstrate that a weak version of this type of reducibility is enough to characterize TTC.

[13]Specifically, we consider an assignment problem to be trivial if it is possible to assign every student to her favorite school. TTC contains a dominant subgroup whenever this is not possible.

Definition 4. Consider an assignment problem (I, R, q) and any capacity vector $0 < q^* < q$. Let $A = \{i \in I | \phi(I, R, q^*)(i) \neq \emptyset\}$ and $UA = \{i \in I | \phi(I, R, q^*)(i) = \emptyset\}$ be the set of assigned and unassigned students, respectively. Then q^* is a **dominant subcapacity** of (I, R, q) if

$$\phi(I, R_A, R'_{UA}, q) = \phi(I, R, q^*) \wedge \phi(UA, R'_{UA}, q - q^*)$$

for any preferences R'_{UA} of the unassigned agents. We call A the **dominant subgroup** associated with q^*.

Definition 5. A mechanism ϕ is **reducible** if all assignment problems $\Gamma = (I, R, q)$ contain a dominant subcapacity whenever some school has a capacity greater than one.

Efficiency imposes a constraint on dominant subcapacities.

Lemma 3. Suppose ϕ is efficient. Given an assignment problem (I, R, q), let q^* be a dominant subcapacity, and let $A = \{i \in I | \phi(I, R, q^*)(i) \neq \emptyset\}$ be the associated dominant subgroup. Then for every $i \in A$, if $aP_i\phi(I, R, q^*)(i)$ then $q^*_a = q_a$.

Proof. Let $UA = I \setminus A$. Suppose for contradiction that there is an $i \in A$, $\phi(I, R, q^*)(i) = b$, aP_ib, and $q^*_a < q_a$. For each $j \in UA$, let R'_j be any preferences such that bP'_ja. Since q^* is a dominant subcapacity, $\phi(R_A, R'_{UA}, q)(i) = b$. Since $q^*_a < q_a$, either a is not assigned to its full capacity or else there is a $j \in UA$ such that $\phi(R_A, R'_{UA}, q)(j) = a$. In the first case, $\phi(R_A, R'_{UA}, q)$ can be Pareto improved by assigning i to a. In the second case, $\phi(R_A, R'_{UA}, q)$ can be Pareto improved by assigning i to a and j to b. Either contradicts the efficiency of ϕ. $\square$

Theorem 2. TTC is the only assignment mechanism that is strategyproof, efficient, just, and reducible.

Proof. Fix an I, R, O, and $\succ$. Consider any strategyproof, efficient, just, and reducible mechanism ϕ. We prove $\phi = TTC$ by induction. Theorem 1 establishes the base inductive step. Consider a capacity vector q such that there is an a with $q_a > 1$. Our inductive hypothesis is that for all $q' < q$, $\phi(I, R, q') = TTC(I, R, q')$. Since ϕ and $q_a > 1$, (R, q) contains a dominant subcapacity q^*. Let $A = \{i \in I | \phi(I, R, q^*)(i) \neq \emptyset\}$ and $UA = \{i \in I | \phi(I, R, q^*)(i) = \emptyset\}$. Therefore, $\phi(I, R, q) = \phi(I, R, q^*) \wedge \phi(UA, R_{UA}, q - q^*)$. By the inductive hypothesis, $\phi(I, R, q^*) = TTC(I, R, q^*)$ and $\phi(UA, R_{UA}, q-q^*) = TTC(UA, R_{UA}, q-q^*)$.

In general, $TTC(I, R, q)(i) \neq TTC(I, R, q')(i)$. However, we show that for a dominant subcapacity, the two coincide. For $TTC(R, q^*)$, we process the cycles one at a time. Fix any ordering of the cycles and consider the first cycle processed, $C_1 = \{o_1, i_1, o_2, i_2, \ldots, o_n, i_n\}$. The set of agents are identical under $TTC(I, R, q^*)$ and $TTC(I, R, q)$, so each object with available capacity under q^* points to the same agent under q^* or q. The agents potentially have a larger set of objects to point at under q than q^*. Therefore, if any agent points to school a under q and b under q^*, it must be that $q_a > q_a^*$ and aP_ib. Therefore, it must be that each student $i_k \in C$ points to o_{k+1} in $TTC(R, q)$; otherwise, if i_k points to a different school a, $q_a > q_a^*$ and $aP_iTTC(R, q^*)(i)$, contradicting Lemma 3. Therefore, C_1 is a cycle both in $TTC(R, q^*)$ and $TTC(R, q)$. After removing it from both, the same logic implies that the second cycle in $TTC(R, q^*)$ is also a cycle in $TTC(R, q)$, and so on. Every time we process a cycle in $TTC(R, q^*)$, we can process the same cycle in $TTC(R, q)$. Therefore, $TTC(I, R, q)(i) = TTC(I, R, q^*)(i)$. Even stronger, when there are no schools with available capacity in $TTC(R, q^*)$, $TTC(UA, R_UA, q - q^*)$ exactly corresponds to the reduced problem in $TTC(I, R, q)$ after removing the same cycles.

Therefore, ϕ and TTC make the same assignments.

$\square$

Clinch and Trade is strategyproof, efficient, and just. Since it does not make the same assignments as TTC, Theorem 2 implies it is not reducible.[14] A serial dictatorship is strategyproof, efficient, and reducible; however, it is not just. Example 3 provides a trivial algorithm that is strategyproof, just, and reducible but not efficient. Example 4 provides an algorithm that is efficient, just, and reducible but is not strategyproof. Therefore, the conditions are independent.

Example 3. Consider the following trivial variation of TTC. We run TTC unless we *start* with two agents, $\{i, j\}$; two objects, $\{a, b\}$; each object has a capacity of one; and the objects have the following priorities:

$\succ_a$	$\succ_b$
i	j
j	i

In this case, we assign i to a and j to b regardless of the preferences they submit. This is clearly strategyproof as preferences are disregarded. It is just since their is no justified envy. It is trivially reducible. However, it is not efficient in the case that i prefers b to a and j prefers a to b.

Example 4. Consider the following variation of TTC. We run TTC unless we start

[14] Under Clinch and Trade, a top trading cycle no longer forms a dominant subproblem. This occurs when a student not in the cycle changes her preferences so that she clinches an object, and her clinch causes one of the agents in the cycle to clinch. It is easy to design an example where this changes the assignment of one of the other agents in the cycle.

with the following problem (here each object has a capacity of one).

R_i	R_j	R_k	$\succ_a$	$\succ_b$	$\succ_c$
b	a	a	i	j	i
c	b	b	k	k	i
a	c	c	j	i	k

In this case, we assign i to c, j to b, k to a, and leave i unassigned. This mechanism is reducible as reducibility has no bite when objects have a capacity of one (and otherwise we run TTC which is reducible). The mechanism is efficient, and it is just since in the only case where we deviate from TTC, we make an assignment with no justified envy. However, it is not strategyproof. If i reports $R'_i : b, a, c$, then we run TTC and i is assigned to b instead of c.

4 Conclusion

It is well known that eliminating justified envy is inconsistent with making a Pareto efficient assignment. This paper introduces an alternative fairness notion, justness, and demonstrates that it is possible to make just and efficient assignments with a strategyproof mechanism. In particular, TTC is the unique mechanism that is strategyproof, efficient, and just.

A reasonable way to define a mechanism Φ as being *fairer* than a mechanism Ψ is if the instances of justified envy for Φ are a subset of the instances of justified envy for Ψ. Our characterization demonstrates that when objects have capacity of one, there is no strategyproof and efficient mechanism that induces strictly fewer instances of justified envy than TTC. Under this interpretation, although no mechanism can be strategyproof, efficient and fair, there is no strategyproof and efficient mechanism that is fairer than TTC.

References

ABDULKADIROĞLU, A., AND Y. CHE (2010): "The Role of Priorities in Assigning Indivisible Objects: A Characterization of Top Trading Cycles," *Mimeo*.

ABDULKADIROĞLU, A., P. PATHAK, AND A. ROTH (2005): "The New York high school match," *The American Economic Review*, 95, 364–367.

——— (2009): "Strategyproofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *The American Economic Review*, 99, 1954–1978.

ABDULKADIROĞLU, A., AND T. SÖNMEZ (1999): "House allocation with existing tenants," *Journal of Economic Theory*, 88(2), 233–260.

ABDULKADIROĞLU, A., AND T. SÖNMEZ (2003): "School choice: A mechanism design approach," *The American Economic Review*, 93(3), 729–747.

ABDULKADIROĞLU, A. PATHAK, P. R. A., AND T. SÖNMEZ (2005): "The Boston public school match," *The American Economic Review*, 95, 368–371.

BALINSKI, M., AND T. SÖNMEZ (1999): "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 84(1), 73–94.

BOGOMOLNAIA, A., AND H. MOULIN (2001): "A new solution to the random assignment problem," *Journal of Economic Theory*, 100(2), 295–328.

CHE, Y.-K., AND F. KOJIMA (2010): "Asymptotic equivalence of probabilistic serial and random priority mechanisms," *Econometrica*, 78(5), 1625–1672.

DUR, U. (2013): “A Characterization of the Top Trading Cycles Mechanism for the School Choice Problem,” *mimeo*.

GALE, D., AND L. SHAPLEY (1962): “College admissions and the stability of marriage,” *American Mathematical Monthly*, 69(1), 9–15.

KESTEN, O. (2010): “School choice with consent,” *The Quarterly Journal of Economics*, 125(3), 1297–1348.

KOJIMA, F., AND M. MANEA (2010a): “Axioms for deferred acceptance,” *Econometrica*, 78(2), 633–653.

KOJIMA, F., AND M. MANEA (2010b): “Incentives in the probabilistic serial mechanism,” *Journal of Economic Theory*, 145(1), 106–123.

KOJIMA, F., AND U. ÜNVER (forthcoming): “The “Boston” School-Choice Mechanism: An Axiomatic Approach,” *Economic Theory*.

LIU, Q., AND M. PYCIA (2013): “Ordinal efficiency, Fairness, and Incentives in Large Markets,” *working paper*.

MORRILL, T. (2013a): “An Alternative Characterization of the Deferred Acceptance Algorithm,” *International Journal of Game Theory*, pp. 19–28.

——— (2013b): “An Alternative Characterization of Top Trading Cycles,” *Economic Theory*, 54, 181–197.

MORRILL, T. (2014a): “Making Efficient School Assignment Fairer,” *working paper*.

MORRILL, T. (2014b): “Two Simple Variations of Top Trading Cycles,” *Economic Theory*.

PÁPAI, S. (2000): “Strategyproof assignment by hierarchical exchange,” *Econometrica*, 68(6), 1403–1433.

PATHAK, P. A., AND T. SÖNMEZ (2012): “School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation,” *The American Economic Review*.

PYCIA, M., AND U. ÜNVER (2014): “Incentive Compatible Allocation and Exchange of Discrete Resources,” *mimeo*.

ROTH, A. (1982): “The economics of matching: Stability and incentives,” *Mathematics of Operations Research*, pp. 617–628.

ROTH, A., AND M. SOTOMAYOR (1992): *Two-sided matching: A study in game-theoretic modeling and analysis*. Cambridge Univ Pr.

SÖNMEZ, T., AND U. ÜNVER (2010): “House Allocation with Existing Tenants: A Characterization,” *Games and Economic Behavior*, 69(2), 425–445.