An Alternative Characterization of Top Trading Cycles

Thayer Morrill†

December 2011

Abstract

This paper introduces two new characterizations of the Top Trading Cycles algorithm. The key to our characterizations is a new condition, *independence of irrelevant rankings (IIR)*. Intuitively, a mechanism satisfies IIR if whenever an agent's ranking at an object is irrelevant to her assignment, then it is irrelevant to the assignment of all agents. We demonstrate that a mechanism is Pareto efficient, strategy proof, IIR, and satisfies mutual best if and only if it is Top Trading Cycles. This provides a new insight into what distinguishes Top Trading Cycles from all other efficient and strategy proof assignment mechanisms. We provide a second characterization in terms of

†North Carolina State University. Email address: thayer_morrill@ncsu.edu. I would like to thank Melinda Morrill and Robert Hammond for their helpful comments. I am grateful to an anonymous referee and editor who noticed an error in a previous version of this paper. This paper changed substantially as a result of their comments.

weak Maskin monotonicity. A mechanism satisfies Pareto efficiency, weak Maskin monotonicity, IIR, and mutual best if and only if it is Top Trading Cycles. This allows us to directly compare Top Trading Cycles to known characterizations of the Deferred Acceptance algorithm in terms of weak Maskin monotonicity.
Key Words: Top Trading Cycles, School Choice, Assignment.
JEL Classification: C78, D61, D78, I20

1 Introduction

School choice assignment mechanisms have been a recent contribution of economics to practical market design. In their groundbreaking paper, Abdulkadiroglu and Sonmez (2003) describe the trade-offs between two classic assignment algorithms: Gale's Top Trading Cycles algorithm (hereafter, TTC) and Gale and Shapley's Deferred Acceptance algorithm (hereafter, DA). While both mechanisms are strategy proof, TTC is Pareto efficient but not necessarily stable. Alternatively, DA is stable, but not necessarily efficient. Therefore, while implementing either assignment algorithm is straightforward, the choice of which algorithm to use is difficult.

A policy maker who advocates TTC should be able to answer the following two questions: why use TTC instead of using DA, and why use TTC instead of using an alternative Pareto efficient, strategy proof mechanism?[1] While the answer to either question involves trade-offs, we seek to inform such an

[1]For example, a *serial dictatorship* is efficient and strategy proof. In a serial dictatorship students select in sequence their most preferred school among those that still have available capacity.

analysis by providing two new characterizations of TTC. The first characterization distinguishes TTC from all other Pareto efficient, strategy proof mechanisms. The second allows for a direct comparison between TTC and a known characterization of DA.

We demonstrate that two characteristics of TTC distinguish it from all other strategy proof and efficient mechanisms. The first condition, *mutual best* (MB), specifies that a mechanism satisfy a minimal level of fairness. A mechanism satisfies MB if a student with highest priority at her most preferred school is always assigned that school. Note that a serial dictatorship does not satisfy MB. MB is necessary for stability, but it is a far weaker condition.

To complete the characterization, we introduce a new property, independence of irrelevant rankings. We define a set of agents to be *invariant* if their assignments do not depend on the preferences or rankings of others. This is motivated by MB. If a mechanism satisfies MB, then an agent is always assigned her most preferred object if she has highest priority for the object regardless of her ranking at any other object or the preferences of the other agents. For an invariant set of agents, their rankings at outside objects are irrelevant to their assignment. *Independence of irrelevant rankings* (IIR) specifies that the assignment of every agent is independent of these irrelevant rankings.

IIR is a counterpart to nonbossiness. A mechanism is nonbossy if any change in an agent's preferences that does not affect her assignment does not affect the assignment of other agents. An analogous condition for priorities might be that if a change in an agent's priority for an object does not affect her assignment, then it does not affect the assignment of other agents.

Such a condition would not be satisfied by either TTC or DA.[2] However, IIR is a significantly weaker condition precisely because invariance is such a strong condition. IIR only imposes a restriction on invariant sets. Indeed, we demonstrate that TTC, DA, the Boston mechanism, and random serial dictatorships all satisfy IIR.

Our main characterization distinguishes TTC from all other strategy proof and efficient mechanisms. Specifically, we demonstrate that TTC is the unique mechanism that is strategy proof, efficient, IIR, and satisfies MB. IIR and MB are all intuitive and desirable attributes of a mechanism. Therefore, this characterization provides a positive answer to the motivating question.

We introduce a second characterization in order to compare TTC to DA. We prove that a mechanism satisfies MB, weak Maskin monotonicity, IIR, and efficiency if and only if it is TTC. R' is a *monotonic transformation* of R at μ if for all agents, any object that is preferred to μ under R' is preferred to μ under R. An assignment mechanism ϕ satisfies *weak Maskin monotonicity* if every agent weakly prefers $\phi(R')$ to $\phi(R)$ whenever R' is a monotonic transformation of R at $\phi(R)$. This characterization was motivated by Kojima and Manea (2010) who provide a characterization of DA in terms of non-wastefulness,[3] weak Maskin monotonicity, and population monotonicity.[4] Indeed, the advantage of our characterization is we may make a direct

[2]This is the motivation for the efficiency-adjusted deferred acceptance algorithm introduced in Kesten (2010). In this modification of deferred acceptance, a student may consent to waive her priority at an object if this priority does not affect her assignment but results in a Pareto inefficient assignment for the other agents.

[3]An assignment rule is non-wasteful if for every preference profile, any object that an agent prefers to her assignment has been allocated up to its quota to other agents.

[4]Intuitively, an allocation rule is population monotonic if agents are weakly better off

comparison between TTC and DA. Morrill (2011) demonstrates that a mechanism satisfies MB, non-wastefulness, weak Maskin monotonicity, and population monotonicity if and only if it is DA. Non-wastefulness is only necessary to rule out degenerate assignment mechanisms and is satisfied by TTC. Similarly, DA satisfies IIR. Therefore, one way of distinguishing between TTC and DA is that of efficiency versus population monotonicity.

Throughout the paper, we assume that objects may be assigned to at most one agent. We demonstrate in Section 4 that our characterizations do not hold when objects may be assigned to more than one agent. In particular, when objects may have capacities greater than one, TTC violates IIR. This is surprising as DA, the Boston mechanism, and random serial dictatorships all continue to satisfy IIR for arbitrary capacities. It remains an open question how to characterize TTC when objects may be assigned to more than one agent.

1.1 Relation to the Literature

This paper is closely related to Abdulkadiroglu and Che (2010), hereafter *A&C*. Their paper is the first to ask what distinguishes top trading cycles from all other strategy proof and efficient mechanisms. They provide an answer by introducing the concept of *recursively respecting top priorities*. They define an assignment to *respect top priorities* if for every student i, the student that has top priority at i's assignment prefers her own assignment to i's assignment. If an assignment respects top priorities, then consider any

when fewer other agents participate in the assignment. A formal definition is provided on Page 21.

student i_1 and label i_1's assignment s_1. Let i_2 be the student with highest priority at s_1. Since the assignment respects top priorities, i_2's assignment s_2 must be preferred by i_2 to s_1. Similarly, label the student with highest priority at s_2 i_3, and so on. Since there are only a finite number of students, eventually we must repeat a student. *A&C* call this a *top priority group*. They define a matching to *recursively respect top priorities* if it respects top priorities and every time a top priority group and their assignments are removed, the matching on the sub-population continues to respect top priorities. *A&C* demonstrate that the unique mechanism that is strategy proof, efficient, and recursively respects top priorities is TTC.

Our paper is intended to complement *A&C*'s characterization. As *A&C* point out, a full characterization of TTC is important for a school board deciding how to best implement school choice. Our characterizations give alternative ways of explaining TTC.

Mutual best is logically independent of both respecting top priorities and recursively respecting top priorities. Notice that the null assignment respects top priorities, but it is straightforward to find a preference profile where the null assignment violates MB. However, for efficient mechanisms, MB is a strictly weaker concept than respecting top priorities. Specifically, if ϕ is non-wasteful and respects top priorities, then ϕ satisfies MB.[5] The example below demonstrates that MB does not imply respecting top priorities. Therefore,

[5]The argument is as follows. Consider a non-wasteful mechanism ϕ that respects top priorities, and let $i \in I$, $a \in O$, $R \in \mathcal{R}$, and $\succ \in \mathcal{C}$ be such that aP_ib for all $b \in O \setminus \{a\}$ and $i \succ_a j$ for all $j \in I \setminus \{i\}$. Then $\phi(a) \neq \emptyset$ or else ϕ would be wasteful. If $\phi(i) \neq a$ and $\phi(a) \neq \emptyset$, then for some $j \in I \setminus \{i\}$, $\phi(j) = a$. Therefore, i has top priority at j's assignment, but i prefers j's assignment to her own. This violates respecting top priorities. Therefore, $\phi(i) = a$ and ϕ satisfies MB.

MB is a strictly weaker assumption than respecting top priorities for non-wasteful assignment mechanisms.

Example 1. Consider the following mechanism ϕ. When there are three students, i, j, and k, three objects, a, b, and c, and the agents and objects have the following preferences and priorities, then ϕ assign i, j, and k to c, a, and b, respectively. For any other assignment problem, ϕ runs TTC.

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

ϕ satisfies MB since in the special case, no agent has highest priority at her top choice and TTC satisfies mutual best. However, ϕ does not respect top priorities as i has top priority at a and i prefers a to c. Of course, ϕ also does not recursively respect top priorities.

Our paper is also closely related to Kojima and Manea (2010) and Morrill (2011). Kojima and Manea (2010) provide two characterizations of when an allocation rule corresponds to the deferred acceptance algorithm for some substitutable priorities of the objects being assigned. Morrill (2011) characterizes when a mechanism is equivalent to the deferred acceptance algorithm for all substitutable priorities. Our first characterization was deliberately chosen to be in terms of weak Maskin monotonicity so that we could directly compare top trading cycles to this characterization of the deferred acceptance algorithm. Ehlers and Klaus (2010) provide a related characterization of DA when priority rules are limited to be responsive.

There are a number of other papers characterizing assignment mechanisms.

Papai (2000) characterizes hierarchical exchange rules, which top trading cycles is a particular instance of, in terms of Pareto efficiency, group strategy proofness, and reallocation proofness. Pycia and Unver (2010) introduce and characterize a generalization of top trading cycles they call trading cycles with brokers and owners. There have been several recent papers characterizing assignment mechanisms other than top trading cycles. Kojima and Unver (2010) provide a characterization of the Boston assignment mechanism.

2 Model

We consider a finite set of agents $I = \{1, \ldots, n\}$ and a finite set of objects $O = \{a, b, c, \ldots\}$. We assume that each object may be assigned to at most one agent. We demonstrate in Section 4 that our characterizations do not extend when objects may be assigned to more than one agent. Each agent $i \in I$ has a complete, irreflexive, and transitive preference relation P_i over $A \cup \{\emptyset\}$. $\emptyset$ represents an agent being unassigned, and $q_\emptyset = \infty$. aP_ib indicates 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 $a = b$.

Each object $a \in O$ has a complete, irreflexive, and transitive priority rule $\succ_a$ over I. In particular, $i \succ_a j$ is interpreted as agent i has 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 and O 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 $a \in O$, $|\{i \in I|\mu(i) = a\}| \leq 1$.[6] In a slight abuse of notation, for a set of agents $I' \subset I$, we define $\mu(I') = \{a \in O|\exists i \in I'$ 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 assignments, respectively. An **allocation mechanism** is a function $\phi : \mathcal{R} \times \mathcal{C} \to \mathcal{A}$. A mechanism ϕ is **strategy proof** if reporting true preferences is each agent's dominant strategy. That is:

$$\phi(P, \succ)(i) R_i \phi(P'_i, P_{-i}, \succ)(i)$$

for all P, $\succ$, $i \in I$, and P'_i.

Abdulkadiroglu and Sonmez (2003) give detailed descriptions of TTC, DA, and the Boston mechanism. Although we characterize TTC when schools have a capacity for only one student, we describe the general algorithms here. 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

[6] In the general case, for each $a \in O$, $|\{i \in I|\mu(i) = a\}| \leq q_a$ where q_a is the capacity of a.

is removed, and the capacity of o_{j+1} is reduced by one. When a school has no remaining capacity, it is removed. For any $R \in \mathcal{R}$, $\succ \in \mathcal{C}$, the mechanism $TTC(R, \succ)$ outputs the assignment 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 (1990) is an excellent resource for the properties of DA.

A third assignment mechanism we consider is the Boston mechanism. In the first round of the Boston mechanism, each student applies to her highest ranked school, and each school accepts up to its capacity the students with highest priority among those that have applied. Each student who was accepted and those schools that are at capacity are removed. In the second round, each remaining student applies to her most preferred school among those that have available capacity. Each school accepts up to its capacity the highest priority students among those that applied. Again, the accepted students and the schools that are at capacity are removed. The process continues until all students are assigned.

3 Alternative Characterizations of Top Trading Cycles

In this section we provide several characterizations of TTC when objects may be assigned to at most one agent. Our two main characterizations are in terms of strategy proofness and weak Maskin monotonicity. We also provide a characterization in terms of Maskin monotonicity. In the Appendix, we provide examples that demonstrate the independence of the conditions used in the characterizations.

3.1 Strategy proofness

Our first characterization of TTC is in terms of strategy proofness and efficiency. This provides a second answer to the following question posed by *A&C*: what distinguishes TTC from all other strategy proof and efficient mechanisms? TTC is the only efficient and strategy proof mechanism that satisfies mutual best and independence of irrelevant rankings. Perhaps what is most surprising about our result is that TTC is characterized without any recursive properties.

Mutual best is a necessary condition for a mechanism to be stable; however, it is a far weaker condition. Stability is typically interpreted as a fairness condition for an assignment mechanism. We interpret mutual best as a mechanism satisfying a minimal level of fairness.

Definition 1. A mechanism ϕ satisfies **mutual best (MB)** if for every $i \in I$ and $a \in O$ such that

- aP_ib for every $b \in O \setminus \{a\}$
- $i \succ_a j$ for every $j \in I \setminus \{i\}$

then $\phi(R, \succ)(i) = a$.

This paper introduces two new properties: invariance and independence of irrelevant rankings. Intuitively, a set of agents is invariant if their assignment does not depend on the preferences or rankings of others. For an invariant set of agents, their rankings at objects not assigned to members of the set are irrelevant to their assignment. Independence of irrelevant rankings specifies that the assignment of any agent is independent of these irrelevant rankings.

Below, we formalize the intuitive notion of adjusting the rank of an agent at a given object.

Definition 2. Given $I' \subset I$ and $O' \subset O$, $\succ_{O'}$ is an I'**-ranking adjustment** of $\succ'_{O'}$ if for every $k, l \in I \setminus I'$ and every $a \in O'$

$$k \succ_a l \Leftrightarrow k \succ'_a l.$$

In words, a ranking adjustment changes the priorities of a set of agents at an object but leaves the relative rankings of the other agents unchanged. Note that being a ranking adjustment is a reciprocal relationship.

Definition 3. Given an assignment $\phi(R, \succ) = \mu$, a set $I' \subset I$ is $\phi(R, \succ)$**-invariant** if for every I'-ranking adjustment $\succ'_{O \setminus \mu(I')}$ of $\succ_{O \setminus \mu(I')}$ and every $R'_{I \setminus I'}$:

$$\phi(R, \succ)(i) = \phi(R_{I'}, R'_{I \setminus I'}, \succ_{\mu(I')}, \succ'_{O \setminus \mu(I')})(i), \text{ for all } i \in I'.$$

Although invariance has a technical definition, it is an intuitive concept. A set of agents is invariant if its members assignments do not depend on their ranks at outside objects or the preferences of agents not in the set. This is motivated by MB. For a mechanism that satisfies MB, an agent with highest priority at her most preferred object is an invariant set. She is always assigned that object regardless of her priority at any other object or the preferences of any other agent.

Invariance is deliberately defined to be a very strong condition. However, for any $\phi(R, \succ)$ there always exists an invariant set as I is trivially invariant. An invariant set's priorities at other objects are irrelevant to the agents in the set. Independence of irrelevant rankings specifies that these priorities are irrelevant to all agents.

Definition 4. A mechanism ϕ is **independent of irrelevant rankings (IIR)** if for every $R \in \mathcal{R}$, $C \in \mathcal{C}$, and every $\phi(R, \succ)$-invariant $I' \subset I$, then for every I'-ranking adjustment $\succ'_{O \setminus \mu(I')}$:

$$\phi(R, \succ)(i) = \phi(R, \succ_{\mu(I')}, \succ'_{O \setminus \mu(I')})(i), \text{ for all } i \in I$$

where $\mu = \phi(R, \succ)$.

By definition, an I'-ranking adjustment does not affect the assignment of any member of I' when I' is an invariant set. IIR specifies that this adjustment which is irrelevant to I' is irrelevant to the other agents as well. This is a desirable property in terms of fairness.

Invariance is a very strong condition on a set. It is precisely the strength of invariance that makes IIR a mild assumption. As the next lemma demonstrates, the standard assignment algorithms all satisfy IIR.

Lemma 1. Top trading cycles satisfies IIR when objects may be assigned to at most one agent. The deferred acceptance algorithm, the Boston mechanism, and a random serial dictatorship all satisfy IIR regardless of the capacities of the objects being assigned.

The proof is in the appendix. The basic intuition is that invariance is such a strong condition that it is satisfied only by very specialized sets. For example, cycles in the first round of TTC are invariant as they are independent of the other agents' preferences or the rankings of the other objects. Similarly, cycles from the first round of TTC plus any cycles from the second round form an invariant set. We show that a chain of cycles defined by TTC are the only invariant sets for TTC. By construction, these agents' priorities at other objects are irrelevant to the assignment made by TTC.

For an example of a mechanism that violates IIR, consider the following assignment rule. Fix an agent i_1. i_1 is the dictator and may choose any object she wishes. Fix two agents $\{i_2, i_3\} \subset I \setminus \{i_1\}$. If i_1 has highest priority at every object other than her top choice, then i_2 is the next dictator. Otherwise, i_3 is the next dictator. The remaining dictators can be chosen arbitrarily. The set $\{i_1\}$ is $\phi(R, \succ)$-invariant for any R and $\succ$ as i_1 is always assigned her top choice. However, this mechanism violates IIR as i_1's ranking at the other objects determines which agent is the next dictator. Lemma 1 demonstrates that a random serial dictatorship satisfies IIR.

Lemma 2 provides the base inductive step for the proof of the strategy proof characterization. However, it is also of independent interest.

Lemma 2. Let ϕ be any mechanism that is strategy proof, efficient, and satisfies mutual best. Let $I_1(R, \succ)$ be the agents assigned in the first round of $TTC(R, \succ)$.[7] Then:

1. $\phi(R, \succ)(i) = TTC(R, \succ)(i)$ for every $i \in I_1(R, \succ)$.

2. $I_1(R, \succ)$ is $\phi(R, \succ)$-invariant.

Proof. Suppose for contradiction there exists a $R \in \mathcal{R}$, $\succ \in \mathcal{C}$, and $i \in I_1(R, \succ)$ such that $\phi(R, \succ)(i) \neq TTC(R, \succ)(i)$. For notational convenience, we will fix $\succ$ for the remainder of the proof. Let $C = \{o_1, i_1, o_2, i_2, \ldots, o_K, i_K\}$ be i's cycle in the first round of $TTC(R)$ where i_j has highest priority at o_j and o_{j+1} is i_j's most preferred object (where it is understood that o_{K+1} refers to o_1). Without loss of generality, let $i = i_K$.

For $i_j \in C$ define $R^1_{i_j} := o_{j+1}, o_j, \emptyset$ and define $R^2_{i_j} := o_j, \emptyset$. Note that for any $i_j \in C$ and any preferences R'_{-i_j} of the other agents, MB implies that $\phi(R^2_{i_j}, R'_{-i_j})(i_j) = o_j$ as i_j has highest priority at o_j. Therefore, for any $R' \in \mathcal{R}$ and any $i_j \in C$, $\phi(R')(i_j) R'_{i_j} o_j$ or else i_j could strictly improve her assignment by reporting $R^2_{i_j}$ which would violate strategy proofness. Therefore, for each $i_j \in C$ and any preferences R'_{-i_j},

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

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, Eq. (1) implies that $\phi(R^1_{i_K}, R_{-i_K})(i_{K-1}) = o_{K-1}$. Therefore, $\phi(R^1_{i_{K-1}}, R^1_{i_K}, R_{-\{i_{K-1}, i_K\}})(i_{K-1}) = o_{K-1}$ or else i_{K-1} could

[7] We have defined TTC to assign all cycles that are present. However, Lemma 2 holds if we assign the cycles one at a time or if we assign any subset of the cycles in the first round of TTC.

profitably misreport her preferences. 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 $i \in I_1(R, \succ)$.

In the above argument, the preferences of agents in $I \setminus I_1$ and the priorities of objects in $O \setminus O_1$ are irrelevant. Therefore, I_1 is $\phi(R, \succ)$-invariant. □

Lemma 2 tells us that if we want a mechanism to be strategy proof, efficient, and satisfy a minimal level of fairness (MB), then we must accept the violations of fairness caused by trades in the first round of TTC. Theorem 1 demonstrates that if we want a mechanism to also satisfy a minimal level of consistency (IIR), then the fairness distortions caused by TTC are necessary. It was well known that it is impossible for a mechanism to be strategy proof, efficient, and stable.[8] However, our results show that TTC can be interpreted as the most fair among strategy proof, efficient mechanisms that are minimally consistent.

[8]Roth (1982) points out that a stable and efficient assignment need not always exist. Kesten (2010) demonstrates that there is no efficient and strategy proof mechanism that always selects an efficient and stable assignment when it exists.

Theorem 1 (*strategy proof Characterization*)**.** A mechanism ϕ is strategy proof, efficient, IIR, and satisfies mutual best if and only if $\phi(R, \succ) = TTC(R, \succ)$ for all $R \in \mathcal{R}$ and $\succ \in \mathcal{C}$.

Proof. Lemma 1 establishes that TTC satisfies IIR. It is well known that TTC is strategy proof, efficient, and satisfies MB. We state this formally in Lemma 3. For the only if direction, consider any R and $\succ$. Let I_k be the set of agents who are matched in step k of $TTC(R, \succ)$. For convenience, define $I_{<k} := \cup_{1 \leq j < k} I_j$. We proceed by induction on k. For our inductive hypothesis, we assume that $\phi(R, \succ)(i) = TTC(R, \succ)(i)$ for all $i \in I_{<k}$ and that $I_{<k}$ is $\phi(R, \succ)$-invariant. We will prove that $\phi(R, \succ)(i) = TTC(R, \succ)(i)$ for all $i \in I_{<k+1}$ and that $I_{<k+1}$ is $\phi(R, \succ)$-invariant. Lemma 2 establishes the base inductive step. Let $O_{<k} = \{a \in O | \phi(R, \succ)(i) = a \text{ for some } i \in I_{<k}\}$. Suppose for contradiction that there exists an $i \in I_k$ such that $\phi(R, \succ)(i) \neq TTC(R, \succ)(i)$ and let $C = \{o_1, i_1, \ldots, o_K, i_K\}$ be the cycle defined by TTC such that $i \in C$. Without loss of generality, let $i = i_K$.

For $a \notin O_{<k}$, define a priority ordering $\succ'_a$ as follows:

$i \in I_{<k},\ j \notin I_{<k}$	$j \succ'_a i$
$i, j \in I_{<k}$	$i \succ'_a j \Leftrightarrow i \succ_a j$
$i, j \notin I_{<k}$	$i \succ'_a j \Leftrightarrow i \succ_a j$

In words, for objects not assigned to members of $I_{<k}$, $\succ'_a$ lowers the priorities of all agents in $I_{<k}$ below those not in $I_{<k}$. Otherwise, it leaves the relative ordering of agents unchanged. For all $a \in O_{<k}$, let $\succ'_a = \succ_a$. In particular, $\succ'$ is an $I_{<k}$-ranking adjustment of $\succ$ for the objects not assigned to members of $I_{<k}$. $I_{<k}$ is $TTC(R, \succ)$-invariant. By the inductive hypothesis, $I_{<k}$ is

$\phi(R, \succ)$-invariant. Therefore, by IIR, it follows that:

$$TTC(R, \succ) = TTC(R, \succ') \tag{3}$$

and

$$\phi(R, \succ) = \phi(R, \succ'). \tag{4}$$

Since $TTC(R, \succ)(i_K) \neq \phi(R, \succ)(i_K)$, Equations (3) and (4) imply $\phi(R, \succ')(i_K) \neq TTC(R, \succ')(i_K)$.

Consider any $o_j \in C$. By the definition of $TTC(R, \succ)$, $i_j \succ_{o_j} i$ for every $i \in I \setminus I_{<k}$. Therefore, by construction, $i_j \succ'_{o_j} i$ for every $i \in I \setminus \{i_j\}$. For $i_j \in C$ define $R^1_{i_j} := o_{j+1}, o_j, \emptyset$ and define $R^2_{i_j} := o_j, \emptyset$. Repeating the argument in Lemma 2, we find that for every $i_j \in C$ and any preferences R'_{-i_j} of the other agents,

$$\phi(R^2_{i_j}, R'_{-i_j}, \succ')(i_j) = o_j. \tag{5}$$

Similarly,

$$\phi(R^1_{i_j}, R'_{-i_j}, \succ')(i_j) \in \{o_{j+1}, o_j\}. \tag{6}$$

and

$$\phi(R^1_C, R_{-C}, \succ')(i_j) = o_j. \tag{7}$$

However, Eq (7) leads to a contradiction. If $\phi(R^1_C, R_{-C}, \succ')(i_j) = o_j$ 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 $\phi(R^1_C, R_{-C}, \succ')$. Therefore, for every $i \in I_k$, $\phi(R, \succ)(i) = TTC(R, \succ)(i)$. Combining this with the inductive hypothesis yields $\phi(R,\succ)(i) = TTC(R,\succ)(i)$ for every $i \in I_{<k+1}$. The preferences of agents in $I \setminus I_{<k+1}$ and the priorities of objects in $O \setminus O_{<k+1}$ are irrelevant to the preceding argument. Therefore, $I_{<k+1}$ is $\phi(R,\succ)$-invariant. □

It is well known that TTC is strategy proof, efficient, and satisfies MB. We state this formally in Lemma 3.

Definition 5. Preferences R'_i **monotonically transform** R_i at a (R'_i m.t. R_i at a) if for all $b \in O \cup \{\emptyset\}$, $bR'_i a$ implies $bR_i a$. R' is a monotonic transformation at an allocation μ (R' m.t. R at μ) if R'_i m.t.R_i at $\mu(i)$ for each $i \in I$. A mechanism ϕ is **Maskin monotonic** if whenever R' m.t. R at $\phi(R,\succ)$ then $\phi(R',\succ) = \phi(R,\succ)$. A mechanism ϕ is **weakly Maskin monotonic** if whenever R' m.t. R at $\phi(R,\succ)$ then $\phi(R',\succ)R'\phi(R,\succ)$.[9]

Lemma 3. TTC is Maskin monotonic, weakly Maskin monotonic, efficient, strategy proof, and satisfies mutual best.

Proof. All properties have been previously demonstrated in the literature. Papai (2000) and Takamiya (2001) demonstrate that TTC is Maskin monotonic. Maskin monotonicity implies weak Maskin monotonicity. Papai (2000) demonstrates that TTC is efficient and strategy proof. If an agent has highest priority at her most preferred object, then the agent and object form a cycle in the first round of TTC. Therefore, TTC satisfies mutual best. □

Corollary 2 (*Maskin characterization*)**.** A mechanism ϕ is Maskin monotonic, efficient, IIR, and satisfies MB if an only if $\phi(R,\succ) = TTC(R,\succ)$ for all $R \in \mathcal{R}$ and $\succ \in \mathcal{C}$.

[9]The relationship between weak Maskin monotonicity, Maskin monotonicity, and DA is discussed extensively in Kojima and Manea (2010).

Proof. Maskin monotonicity is equivalent to group strategy proofness for an allocation rule which implies strategy proofness (Takamiya 2001). □

3.2 Weak Maskin Monotonicity

DA satisfies weak Maskin monotonicity but not Maskin monotonicity.[10] We introduce a characterization of TTC in terms of weak Maskin monotonicity so that we may directly compare TTC to a known characterization of DA.

Theorem 3 (*Weak Maskin Characterization*)**.** A mechanism ϕ satisfies weak Maskin monotonicity, efficiency, IIR, and mutual best if and only if $\phi(R, \succ) = TTC(R, \succ)$ for every $R \in \mathcal{R}$ and $\succ \in \mathcal{C}$.

Proof. We prove that a weakly Maskin monotonic and efficient mechanism is Maskin monotonic. The characterization then follows immediate from Corollary 2. Let R' m.t. R at $\phi(R, \succ)$. By weak Maskin monotonicity, $\phi(R', \succ) R' \phi(R, \succ)$. Therefore, for every i either $\phi(R', \succ)(i) = \phi(R, \succ)(i)$ or $\phi(R', \succ)(i) P'_i \phi(R, \succ)(i)$. However, if $\phi(R', \succ)(i) P'_i \phi(R, \succ)(i)$, then $\phi(R', \succ)(i) P_i \phi(R, \succ)(i)$. Therefore, if $\phi(R', \succ) \neq \phi(R, \succ)$ then $\phi(R', \succ)$ Pareto improves $\phi(R, \succ)$ under preferences R which would contradict the efficiency of ϕ. Therefore, if a mechanism is weakly Maskin monotonic and efficient, then the mechanism is Maskin monotonic. □

We are now able to make a direct comparison between TTC and the deferred acceptance algorithm using a characterization from Morrill (2011).[11] That

[10]See Kojima and Manea (2010).

[11]The characterization in Morrill (2011) is based on an earlier characterization by Kojima and Manea (2010).

characterization uses the property of population monotonicity. Intuitively, an allocation rule is population monotonic if agents are weakly better off when they compete against fewer agents. For convenience, population monotonicity is defined in terms of a subgroup finding all objects unacceptable rather than not participating in the assignment. Specifically, let $R^{\emptyset}$ denote the preference profile that ranks $\emptyset$ (being unassigned) first for every agent.

Definition 6. A mechanism ϕ is **population monotonic** if

$$\phi(R_{I'}, R^{\emptyset}_{I \setminus I'})(i) R_i \phi(R)(i) \; \forall i \in I', \forall I' \subset I, \forall R \in \mathcal{R}.$$

As a reminder, an assignment rule is non-wasteful if any object that an agent prefers to her assignment has been allocated up to its quota to other agents.

Theorem 4 (Theorem 1, Morrill (2011))**.** A mechanism ϕ satisfies non-wastefulness, population monotonicity, weak Maskin monotonicity, and mutual best if and only if it is the deferred acceptance algorithm.

Theorem 3 and 4 allow us to make a direct comparison between TTC and DA. Non-wastefulness is only necessary to rule out trivial mechanisms.[12] In particular, DA satisfies weak Maskin monotonicity, mutual best, and IIR while TTC satisfies weak Maskin monotonicity, non-wastefulness, and mutual best. Traditionally, the difference between TTC and DA is described as efficiency versus stability. While efficiency versus stability remains the most

[12] For example, the mechanism that assigns mutual top choices but leaves all other agents unassigned is trivially population monotonic, weakly Maskin monotonic, and satisfies mutual best. Non-wastefulness is only necessary to rule out these types of degenerate mechanisms.

informative comparison, our characterization provides a new distinction between the two mechanisms. TTC is efficient but not population monotonic while DA is population monotonic but not efficient.

Efficiency is one of the most important goals of a market designer. On the other hand, a market designer does not aim to make current agents worse off when the population being assigned is expanded. Rather, this is the expected byproduct of increased competition for the same resources. In this sense, population monotonicity is not inherently desirable but rather a property that we expect to hold. In fact, we may think of situations where agents benefit from the participation of other agents. For example, consider a school board that would like to implement TTC to make school assignments but is only willing to change the assignment mechanism if enough students participate. Suppose for political reasons the board decides they will only run TTC if at least two thirds of the students submit preferences; otherwise, the board will keep the *status quo* assignment. In this scenario, a student who is unhappy with her current assignment may benefit from the participation of other students if otherwise the number of participating students is below the threshold.

4 Top Trading Cycles with general capacities

Interestingly, TTC loses some of its desirable properties when objects may be assigned to more than one agent. For example, Kesten (2006) demonstrates that under TTC an agent is sometimes made worse off when the capacities of some or all of the objects are increased. Similarly, under TTC agents are sometimes worse off when they compete against fewer students for the same objects.[13] In contrast, under DA students are always made better off by increased capacity and decreased competition.

IIR is a new condition that is satisfied by DA but violated by TTC when objects may be assigned to more than one agent. We demonstrate that TTC violates IIR with Example 2. This demonstrates that the characterizations presented in Theorem 1 and 3 do not generalize to the case when objects may be assigned to more than one agent.

Example 2 (TTC violates IIR)**.** Suppose there are three agents $\{i, j, k\}$ and two objects $\{a, b\}$. a's capacity is two, and b's capacity is one. Define R and $\succ$ according to the following rank-order lists:

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

In the first round of TTC, $\{i, b, j, a\}$ form a cycle. Therefore, $TTC(R, \succ)$ assigns i, j, and k to b, a, and a respectively. Moreover, $\{j\}$ is $TTC(R, \succ)$-invariant as j has one of the two highest priorities at her most preferred object which has a capacity of two. Therefore, TTC will always assign j to a regardless of i and k's preferences or j's priority at b. However, consider the following j-ranking adjustment $\succ'_b := k, i, j$. Now, $TTC(R, \succ')$ assigns i, j, and k to a, a, and b respectively. Therefore, when TTC may have capacities greater than one, TTC violates IIR.

[13] These properties are called resource monotonicity and population monotonicity, respectively. See Kesten (2006) for a formal definition.

Notice that in Example 2, the first trade in $TTC(R, \succ)$ is unnecessary; j does not need to trade with i in order to be assigned a. However, this trade causes a distortion. This trade does not affect j's assignment, but it results in i being assigned to b despite k preferring b to a and having higher priority than i at b. Morrill (2012) demonstrates that we may implement the second assignment with a strategy proof, efficient, and MB mechanism. Therefore, we may no longer interpret TTC as being the strategy proof and efficient algorithm that satisfies MB with the minimum number of distortions to fairness.

It remains an open question as to how to characterize TTC when objects may be assigned to more than one agent. Morrill (2012) introduces a variation on TTC that first checks if an agent has one of the q_a highest priorities at her most preferred object a, where q_a is the capacity of a, before allowing her to point to a. If she is one of the q_a highest priority students, then she is assigned a and not allowed to trade her priority at any other object. Note that this mechanism satisfies $A\&C$'s property of recursively respecting top priorities but does not always produce the same assignment as TTC. Therefore, their characterization does not immediately generalize to arbitrary capacities.

5 Conclusion

Abdulkadiroglu and Che (2011) is an important paper for market design both because it increases our understanding of a key algorithm, top trading cycles, and because it provides a way of explaining to a policy maker what differentiates top trading cycles from all other strategy proof and Pareto efficient allocation mechanisms. Our paper makes a similar contribution by providing several new characterizations of TTC.

In particular, we provide separate characterizations in terms of strategy proofness, Maskin monotonicity, and weak Maskin monotonicity. This allows us to make a direct comparison between top trading cycles and the deferred acceptance algorithm. It also provides another explanation of what differentiates top trading cycles from all other strategy proof and efficient mechanisms. In particular, top trading cycles is the only such mechanism that satisfies mutual best and independence of irrelevant rankings. Both of these properties are desirable attributes of an allocation mechanism. Moreover, they are simple and intuitive to explain to a policy maker.

6 Appendix

Lemma 1 *Top trading cycles satisfies IIR when objects may be assigned to at most one agent. The deferred acceptance algorithm, the Boston mechanism, and a random serial dictatorship all satisfy IIR regardless of the capacities of the objects being assigned.*

Proof. Top trading cycles:

Fix R and $\succ$ and let I' be a $TTC(R, \succ)$-invariant set. Let $\mu = TTC(R, \succ)$. For any $i \in I'$, i's most preferred object a must be contained in $\mu(I')$. Otherwise, if $\succ'_a$ is the $\{i\}$-ranking adjustment that gives i highest priority at a, then $TTC(R, \succ_{-a}, \succ'_a)(i) = a \neq TTC(R, \succ)(i)$, violating invariance. Similarly, if $a \in \mu(I')$ and i is the agent with highest priority at a, then $i \in I'$. Otherwise, if i ranks a first, then $\mu(a)$'s assignment under TTC is changed, violating invariance. Therefore, in the first round of TTC, every agent in I' points to an object in $\mu(I')$ and *vice versa*. This implies that there must be a cycle consisting entirely of agents in I' and objects in $\mu(I')$. Choose one cycle (there may be more than one), and label the agents in the cycle A_1. Consider any $i \in I' \setminus A_1$, and let a be i's most preferred object among $O \setminus \mu(A_1)$. As before, it must be that $a \in \mu(I') \setminus \mu(A_1)$. Otherwise, if $\succ'_a$ is the $\{i\}$-ranking adjustment that gives i highest priority at a, then $TTC(R, \succ_{-a}, \succ'_a)(i) = a \neq TTC(R, \succ)(i)$, violating invariance. Similarly, for any $a \in \mu(I') \setminus \mu(A_1)$, the agent with highest priority at a (other than possibly agents in A_1) must be contained in $I' \setminus A_1$. Again, there must be a cycle, this cycle is contained in $I' \setminus A_1$ and $\mu(I') \setminus \mu(A_1)$, and this is a trading cycle in TTC. By iterating this argument we can partition I' into $\{A_1, A_2, \ldots, A_k\}$ where A_i is a set of agents in a cycle in TTC after removing the agents $A_1 \cup \ldots A_{i-1}$ and their assignments. By following this order of choosing cycles, we can assign all of the agents in I' before we consider any agent outside of I'. Since the order in which TTC processes cycles is irrelevant to the final assignment,[14] the priorities of agents in I' at objects outside of $\mu(I')$ are irrelevant to any agent's assignment, and TTC satisfies IIR. Example 2 on Page 23 demonstrates that when objects may be assigned to more than one agent, TTC violates IIR.

Deferred acceptance algorithm:

Fix a R and $\succ$, let I' be a $DA(R, \succ)$-invariant set, set $DA(R, \succ) = \mu$, and let $O' = \mu(I')$. Consider any $i \in I'$. It must be that $\mu(i) P_i a$ for any $a \notin \mu(I')$. Otherwise, if $\succ'_a$ is the $\{i\}$-ranking adjustment that gives i highest priority at a, then $DA(R, \succ_{-a}, \succ'_a)(i) \neq DA(R, \succ)(i)$ as i is never rejected by a and

[14]See Abdulkadiroglu and Sonmez (1999).

i only proposes to $\mu(i)$ if she is rejected by a. This would contradict the invariance of I'. However, i's priority at any object she finds strictly inferior to $\mu(i)$ is irrelevant to i's assignment or any other agent's assignment as i never proposes to that object. Therefore, if we adjusted i's ranking at any object not assigned to a member of I' and reran DA, no assignment would be changed.

The Boston mechanism:

In the Boston mechanism, an agent's priority at an object is only used as a tie-breaker when more than one agent applies to the same object in the same round. Fix a R and $\succ$, and let $\phi(R, \succ)$ assign students according to the Boston mechanism. Suppose I' is $\phi(R, \succ)$-invariant, and let $O' = \phi(R, \succ)(I')$. It must be that for any $a \in O \setminus O'$, i never applies to a. Otherwise, let $\succ'_a$ be an i-ranking adjustment that gives i highest priority at a. This has no affect on the Boston mechanism until i applies to a, but under $\succ'_a$ i is assigned a. As this changes i's assignment, this is a violation of invariance. Since each $i \in I'$ never applies to any $a \notin O'$, the priorities of I' at objects in $O \setminus O'$ are irrelevant.

Random serial dictatorship:

Priorities are irrelevant to a random serial dictatorship. Therefore it trivially satisfies IIR. □

We prove the independence of our axioms in Theorems 1 and Theorem 3 through a series of examples. Example 3 demonstrates the independence of mutual best.

Example 3 (Mutual Best)**.** A serial dictatorship is efficient, strategy proof, weakly Maskin monotonic, and IIR. A serial dictatorship violates mutual best.

Example 4 demonstrates the independence of weak Maskin monotonicity and strategy proofness in the characterizations. The mechanisms in Examples 4 and 5 are discussed in greater detail in Abdulkadiroglu and Sonmez (2003).

Example 4 (Weakly Maskin Monotonic, Strategy Proof)**.** The Boston mechanism is efficient, IIR, and satisfies mutual best. However, it is not strategy proof, weakly Maskin monotonic, or Maskin monotonic.

Example 5 demonstrates the independence of efficiency.

Example 5 (Efficiency)**.** DA is strategy proof, weakly Maskin monotonic and satisfies MB and IIR. It is well known that DA is not efficient.

Example 6 demonstrates the independence of IIR.

Example 6 (IIR)**.** Consider the following special case. There are three agents $\{i, j, k\}$ and three objects $\{a, b, c\}$. Whenever R and $\succ$ are such that i has highest priority at all objects, a is i's most preferred object, bP_jc, and bP_kc, then we define $\phi(R, \succ)$ as follows. If $j \succ_b k$, then $\phi(R, \succ)$ assigns i, j, and k to a, c, and b respectively. For every other case (agents, object, preferences, and priorities), ϕ runs TTC. This mechanism is efficient. ϕ is strategy proof as TTC is strategy proof and in the special case, if j reports that she prefers c to b, then we run TTC and j will still be assigned c. ϕ also satisfies MB as TTC satisfies MB and in the special case, only i has highest priority at her most preferred object. However, ϕ violates IIR. Consider a R and $\succ$ so that we are in the special case. $\{i\}$ is $\phi(R, \succ)$-invariant as ϕ satisfies MB and i has highest priority at her most preferred object a. However, if we consider the $\{i\}$-ranking adjustment $\succ'_b := j, i, k$, then we are no longer in the special case. Therefore, $\phi(R, \succ'_b, \succ_{-b})(j) = TTC(R, \succ'_b, \succ_{-b})(j) = b$. Since $\phi(R, \succ)(j) = c$, this violates IIR.

References

Abdulkadiroglu, A., and Y. Che (2010): "The Role of Priorities in Assigning Indivisible Objects: A Characterization of Top Trading Cycles," *Mimeo.*

Abdulkadiroglu, A., and T. Sonmez (1998): "Random serial dictatorship and the core from random endowments in house allocation problems," *Econometrica*, 66, 689–701.

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

Ehlers, L., and B. Klaus (2009): "Allocation via Deferred-Acceptance under Responsive Priorities," *mimeo.*

Gale, D., and L. Shapley (1962): "College admissions and the stability of marriage," *American Mathematical Monthly*, 69(1), 9–15.

Kesten, O. (2006): "On two competing mechanisms for priority-based allocation problems," *Journal of Economic Theory*, 127(1), 155–171.

——— (2010): "School choice with consent," *The Quarterly Journal of Economics*, 125(3), 1297.

Kojima, F., and M. Manea (2010): "Axioms for deferred acceptance," *Econometrica*, 78(2), 633–653.

Morrill, T. (2011): "An Alternative Characterization of the Deferred Acceptance Algorithm," *International Journal of Game Theory, Forthcoming*.

——— (2012): "A Simple Variation of Top Trading Cycles," *mimeo*.

Pápai, S. (2000): "Strategyproof assignment by hierarchical exchange," *Econometrica*, 68(6), 1403–1433.

Pycia, M., and U. Unver (2010): "Incentive Compatible Allocation and Exchange of Discrete Resources," *SSRN eLibrary*.

Roth, A. (1991): "A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the United Kingdom," *The American Economic Review*, 81(3), 415–440.

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

Takamiya, K. (2001): "Coalition strategy-proofness and monotonicity in Shapley-Scarf housing markets," *Mathematical Social Sciences*, 41(2), 201–214.