The Role of Priorities in Assigning Indivisible Objects: A Characterization of Top Trading Cycles

Atila Abdulkadiroglu and Yeon-Koo Che

Duke University and Columbia University

October 2010
Preliminary

Abstract

The Top Trading Cycles mechanism emerges as a desirable solution in various market design applications. Yet recommendations are made without any rigorous foundation for the role priorities play in assignment. We explain that role by *recursive individual rationality with respect to a priority structure*. We show that a mechanism is Pareto efficient, strategy-proof and recursively individual rational with respect to a priority structure if and only if it is the Top Trading Cycles mechanism defined by that priority structure.

1 Introduction

Various notable market design applications involve allocation of heterogenous indivisible objects without monetary transfers, such as assigning pupils to public schools in a school choice program and rematching kidney patients with donors when patients have donors with incompatible kidneys. A common feature of such problems is that, objects usually rank individuals in a priority order, which resembles individuals' preference orderings of the objects. *Top Trading Cycles* (TTC) and its variants, the roots of which can be traced back to Gale's celebrated top trading cycles algorithm, emerges as a desirable solution to incorporate such priorities in the allocation process.

A typical problem consists of a finite set of individuals, say students, and a set of objects with finite capacities, say schools. Students rank schools in strict preference order and schools rank students in priority order. For example, a child who lives within a certain distance from a school may get neighborhood priority for that school and any tie in priorities may be broken by a fair lottery.

Given strict preference lists of students and strict priority lists of schools, TTC assigns students to schools via the following algorithm (Abdulkadiroğlu and Sönmez 2003): In each round, every student points to the school she prefers most among the remaining schools, and every remaining school points to the student that has the highest priority at that school among all remaining students. A cycle is an ordered list of schools and students $\{o_1, i_1, ..., o_K, i_K\}$ such that school o_k points to student i_k and student i_k points to school o_{k+1} and $o_{K+1} \equiv o_1$. When such a cycle exists, student i_k is assigned to school o_{k+1}, the capacity of school o_{k+1} is decreased by one, the students in the cycle and schools with no more capacity are removed; the process is repeated with the remaining students and schools.

The resulting assignment is *Pareto efficient*, that is, there is no alternative assignment that improves a student's assignment without harming others'. Furthermore, TTC is *strategy-proof*, i.e. it makes truthful reporting of preferences over schools a dominant strategy for every student in the induced preference revelation game.

However many other mechanisms meet these two requirements. For example, given an ordering of students, a corresponding *serial dictatorship* mechanism determines the assignment as follows: Each student is assigned in the given order to her most preferred school among the remaining ones. Any serial dictatorship mechanism is Pareto efficient and strategy-proof as well.

Given the richness of Pareto-efficient and strategy-proof mechanisms, then a natural question arises: Why TTC? In other words, what is the additional property that uniquely pins down TTC among all efficient and strategy-proof mechanisms? Despite its prevalence in market design applications, this question remains open. Without a satisfactory answer, any recommendation of TTC in such applications would not be well-grounded. We fill in this gap by offering a new characterization of TTC.

The most related work to ours is Pápai (2000). In a model in which each object has unit capacity, Pápai characterizes a wide class of mechanisms, hierarchical exchange rules, by Pareto efficiency, group strategy proofness, which rules out beneficial preference manipulation by groups of individuals, and reallocation-proofness, which rules out manipulation by two individuals via misrepresenting preferences and swapping objects ex post. TTC is a hierarchical exchange rule defined by the priority lists of schools. Yet, none of the properties above makes any explicit reference to any specific priority structure. In fact, one of them is about efficiency and the other two are about avoiding sophisticated manipulation schemes by individuals. Therefore, Pápai's result does not address the question of why those specific priorities are used in TTC.[1]

Any normative explanation to the use of specific priorities in TTC should involve an axiom that makes a reference to the priorities. For example, if a student has the highest priority at some school, we might expect her to be assigned to that school whenever that school is her first choice. TTC satisfies that criterion, but a serial dictatorship mechanism does not, which is also a hierarchical exchange rule. Indeed our new axiom, recursive individual rationality, generalizes that requirement for a given priority structure.

Given a priority structure, an assignment is *individually rational for top students* if every student that is ranked highest at a school receives an assignment that she weakly prefers to that school. An assignment is *recursively individually rational with respect to a priority structure*, if, given the priority structure, it is individually rational for top students and it continues to be individually rational for top students in the reduced problems when the top students are removed with their assignments recursively. A mechanism is recursively individually rational with respect to a priority structure if every assignment it produces is recursively individually rational with respect to the given priority structure. Our main result states that a mechanism is Pareto efficient, strategy proof and recursively individually rational with respect to a priority structure if and only if it is the TTC defined by the priority structure.

Our axiom is a generalization of Ma's (1994) individual rationality axiom. Ma studies the housing market problem (Shapley and Scarf 1974), in which each individual owns a house which she would like to exchange for another one she prefers more. The unique core of the market is found via Gale's top

[1] In a similar vein, Pycia and Ünver (2010) introduce and characterize trading cycles with brokers and owners by Pareto efficiency and group strategy-proofness. Also Kojima and Manea (2010) provide a characterization for Gale-Shapley and Kojima and Unver (2010) provide a characterization for the Boston school choice mechanism, both via a monotonicity condition on preferences.

trading cycles algorithm (Postlewaite and Roth 1977), in which each agent points to the owner of his most preferred house among the remaining houses. An allocation is said to be individually rational for a housing market if every agent is assigned a house that she weakly prefers to her initial endowment. Ma shows that Gale's top trading cycles is characterized by Pareto efficiency, strategy proofness and individual rationality. It is in that sense that our generalized axiom provides a foundation for TTC.

Such foundation is important in policy making. A common interpretation of TTC is that it effectively allows students to trade in their priorities for better schools of their choice. However that interpretation may prove difficult in making a case for TTC. For instance, sibling priority is usually granted on the belief that assigning siblings to the same schools benefits the siblings via spillover effects and sharing experiences and their parents via solving transportation and coordination problems. From the point of view of a policy maker, trading in sibling priority for a better choice may be difficult to justify since the sibling priority may have been instituted for encouraging siblings to go to the same school. In contrast, our characterization states that, when a school district has Pareto efficiency, strategy-proofness and recursive individual rationality –i.e. guaranteeing each student a school as good as the ones she is top ranked by – as three policy goals to meet, the unique mechanism that meets those criteria is the TTC defined by the given priority structure. It is worth noting that there is no reference to any trading of priorities in the three stated goals, therefore TTC is justified not by allowing students to trade in their priorities but by three policy goals none of which require trading of priorities.

Our result provides a clear answer to the question of what role priorities play in the allocation of indivisible objects. Another intuitive role that can be attributed to priorities is a monotonicity relation between priorities and the assignment. Namely, a mechanism *respects improvements in priorities* if whenever a student's standing in priorities improves, her assignment weakly improves. Note that this notion does not make any reference to any specific priority structure, so it cannot be used to pin down a single priority structure to define TTC. However, TTC satisfies this requirement. Therefore, as a corollary of our main result, Pareto efficiency, strategy-proofness and recursive individual rationality also implies respecting improvements in priorities.

We formalize our arguments in the following sections.

2 Model

A problem consists of a finite set of agents $I = \{1, ..., n\}$ and a finite set of objects $O = \{a, b, c, ...\}$. To simplify the exposition, we will assume that each object has a single copy, but the arguments can be easily generalized if objects come in multiple copies, such as schools with multiple seats in school choice. An agent can consume at most one object and an object can be consumed by at most one agent. Each agent $i \in I$ has a complete, irreflexive and transitive binary preference relation P_i over $O \cup \{\varnothing\}$ and $\varnothing$ represents consuming nothing. aP_ib means that i prefers a to b. Each object $a \in O$ ranks agents by a complete, irreflexive and transitive binary priority relation $\succ_a$ over A. $i \succ_a j$ means that i has higher priority at a than j.

Let $P = (P_i)_{i \in I}$, $\succ = (\succ_a)_{a \in O}$, $P_{-I'} = ((P_j)_{j \in I - I'})$ and $\succ_{-O'} = (\succ_b)_{b \in O - O'}$. We fix I and O and refer to a problem by $(P, \succ)$.

For $i \in I$ let R_i be the symmetric extension of P_i, that is, for all $a, b \in O \cup \{\varnothing\}$, if aP_ib then aR_ib, and if $a = b$ then aR_ib and bR_ia. Let the indifference relation I_i denote the symmetric part of R_i. Define $\succsim$ similarly.

A **matching** of agents to objects is a function $\mu : A \longrightarrow O$ such that $\mu(i) \subset O$, $|\mu(i)| \leq 1$ for all $i \in I$.

A matching μ (Pareto) **dominates** another matching ν if $\mu(i)R_i\nu(i)$ for all $i \in I$ and $\mu(i)P_i\nu(i)$ for some $i \in I$.

A matching is **Pareto efficient** if it is not dominated by another matching.

A (deterministic) **mechanism** selects a matching for every problem. A mechanism is efficient if it selects an efficient matching for every problem. If φ is a mechanism, let $\varphi(P; \succ)$ denote the matching selected by φ. A mechanism φ is **strategy-proof** if reporting true preferences is a dominant strategy for every agent in the preference revelation game induced by φ, that is

$$\varphi(P; \succ)(i) R_i \varphi(P'_i, P_{-i}; \succ)(i) \tag{1}$$

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

3 Top Trading Cycles (TTC)

Given a problem $(I, O, P, \succ)$, *TTC* finds the matching via the following algorithm:

In the first round of the algorithm, all students and schools are available. In every round of the algorithm,

- Every available object points to its highest priority agent among all available agents. Every agent points to her most preferred object among all available objects.

- A *cycle* $c = \{o_k, i_k\}_{k=1,...,K}$ is an ordered list of objects and agents such that o_k points to i_k and i_k points to o_{k+1} for every $k = 1, ..., K$, where $o_{K+1} = o_1$.

- For every cycle $c = \{o_k, i_k\}_{k=1,...,K}$, match each agent with the object she points to in that cycle and remove the agent and the object. In that case, we say that i_k *trades* o_k *for* o_{k+1}.

- Repeat the algorithm in the next round until no more agents are matched.

4 Charaterization: Recursive Individual Rationality

Definition 1 *Given $\succ$, a matching μ is **individually rational for top students** if every student that is ranked highest at a school is matched with an alternative that she weakly prefers to that school. μ is **recursively individually rational with respect to** $\succ$, if, given the priority structure, it is individually rational for top students and it continues to be individually rational for top students in the reduced problems when the top students are removed with their assignments recursively.*
Formally, define the following recursively: $T_0 = \varnothing$ and $O_0 = \varnothing$. Given $\{T_k, O_k\}_{k=0,...,K}$, let

$$O_{K+1} = O - \bigcup_{k=1,...,K} \bigcup_{i \in T_k} \mu(i),$$

be the set of remaining objects in the reduced problem,

$$\hat{I}_{K+1} = I - \bigcup_{k=1,...,K} T_k$$

be the set of remaining students in the reduced problem and

$$T_{K+1} = \{i \in \hat{I}_{K+1} : \exists o \in O_{K+1} \text{ s.t. } i \succ_o j \ \forall j \in \hat{I}_{K+1}\}$$

be the set of top students in the reduced problem. Also define $e : I \longrightarrow 2^O$: $o \in e(i)$ if and only if for some k, $i \in T_k$, $o \in O_k$ and $i \succ_o j \ \forall j \in \hat{I}_k$. μ is ***recursively individually rational with respect to*** *$\succ$ if for all $i \in I$, $\mu(i) R_i o$ for all $o \in e(i)$.*

The definition trivially extends for mechanisms.

Definition 2 *A mechanism is φ is recursively individually rational if for every $(P, \succ)$, $\varphi(P, \succ)$ is recursively individually rational with respect to $\succ$.*

Theorem 3 *A mechanism φ is Pareto efficient, strategy-proof and recursively individually rational if and only if $\varphi(P, \succ) = TTC(P, \succ)$ for all $(P, \succ)$.*

We defer the proof to the appendix. Although our new axiom is a generalization of Ma's individual rationality axiom, our proof technique is different than his. In Ma's environment, which is a special case of ours, the proof can be carried out by induction on agents. By the recursive nature of our individual rationality axiom, we utilize the TTC algorithm and give a proof by induction on the rounds in which TTC allocates objects. In particular, we show that, a Pareto efficient, strategy-proof and recursively individual rational mechanism must assign the same objects as TTC does to the individuals who get their assignment in the first round of TTC. Given the induction hypothesis that the mechanism assign the same objects as TTC does to the individuals who get their assignment in the first $k-1$ rounds of TTC, we show that it must assign the same objects as TTC does to the individuals who get their assignment in the kth round of TTC.

5 Monotonicity

Another intuitive role that can be attributed to priorities is a monotonicity relation between priorities and the assignment. Namely, if whenever a student's standing in priorities improves, her assignment is expected to improve. We make this formal below.

Definition 4 *$\succ'$ is an improvement in priorities for $i \in I$ if*

$$\begin{aligned} &\succ' \text{ is not equivalent to } \succ, \\ &i \succ j \Rightarrow i \succ' j \, and \\ &\forall j,k \in I - \{i\} : j \succ' k \Leftrightarrow j \succ k \end{aligned}$$

Definition 5 *A mechanism φ is respects improvements in priorities if for all $(P, \succ)$, $i \in I$*

(i) if $\succ'$ is an improvement for i, then $\varphi(P, \succ')(i) R_i \varphi(P, \succ)(i)$; and

(ii) if $\varphi(P, \succ)(i)$ is not i's first choice, then there exists an improvement $\succ'$ for i such that $\varphi(P, \succ')(i) P_i \varphi(P, \succ)(i)$.

Since this notion does not make any reference to any specific priority structure, it does not pin down a TTC with a particular priority structure from the set of Pareto efficient and strategy-proof mechanisms. However, TTC satisfies this requirement. Therefore, as a corollary of our main result, Pareto efficiency, strategy-proofness and recursive individual rationality also implies respecting improvements in priorities.

Corollary 6 *Pareto efficiency, strategy-proof and recursive individual rationality implies respecting improvements in priorities.*

6 Extensions and Discussion

Priorities and priority-based mechanisms play an essential role in the allocation of indivisible objects when monetary transfers are not allowed. The role of priorities in Gale-Shapley's celebrated student proposing deferred acceptance mechanism (Gale-Shapley) is well understood. Gale-Shapley is characterized as the student optimal stable stable matching mechanism. Likewise, the Boston mechanism is the student optimal mechanism which produces a stable matching according to the preference-adjusted priorities in which a student who rank a school higher in her choice list has higher priority than a student who ranks it lower and they are ranked according to the original priority order otherwise. Our result complements the picture by explaining the role priorities play in TTC.

To simplify the exposition of the ideas, we have assumed that each object comes in single copy When objects have multiple copies, such as schools in school choice, an object is removed in the definition of recursive individual rationality when all of its copies are removed in the recursive process. With this modification to the definition of the axiom, the main results follows without any change.

A second assumption is that objects rank individuals in strict priority order. The main result follows directly after the breaking of ties potentially via some lottery but not before. To see this, suppose that there is only one school with one seat and two students with equal priority at the school. Since no mechanism can give both the students the single available seat, no mechanism can guarantee individual rationality with respect to the weak priority structure.

When ties at school priorities are broken randomly, an interesting monotonicity relation between priorities and the random TTC allocation emerges. Namely, consider an improvement in priorities for a student. Then the student's random TTC allocation under the improved priority structure first order stochastically dominates her random TTC allocation under the original priority structure. This follows from the fact that, for any tie breaking, her assignment weakly improves under the improved priority structure. A characterization of random TTC remains an open question.

References

[1] **Abdulkadiroğlu, Atia, and Tayfun Sönmez.** 2003. "School Choice: A Mechanism Design Approach." *American Economic Review*, 93, 729-747.

[2] **Kojima, Fuhito and Mihai Mane** 1994. "Axioms for Deferred Acceptance." *Econometrica*, forthcoming

[3] **Kojima, Fuhito and Utku Ünver** 1994. "The Boston School-Choice Mechanism." mimeo

[4] **Ma, Jinpeng** 1994. "Strategy-Proofness and the Strict Core in a Market with Indivisibilities." *International Journal of Game Theory*, 23, 75-83

[5] **Pápai, Szilvia** 2000. "Strategyproof Assignment by Hierarchical Exchange." *Econometrica*, Vol. 68, No. 6 (Nov., 2000), pp. 1403-1433.

[6] **Pycia, Marek and Utku Ünver** 2010. "Incentive Compatible Allocation and Exchange of Discrete Resources." *mimeo*

[7] **Postlewaite, A. and A.E. Roth** 19tt. "Weak versus Strong Domination in a Market with Indivisibile Goods." *Journal of Mathematical Economics*, 4, 131-137

[8] **Shapley, LS and H. Scarf** 1974. "On cores and Indivisibility." *Journal of Mathematical Economics*, 1, 23-37

A Proof of Theorem 1

TTC is Pareto efficient and strategy-proof. It is trivial to show that it is also recursively individual rational.

Let φ be a Pareto efficient, strategy-proof and recursively individually rational mechanism. To the contrary, suppose that there exists $(P, \succ)$ such that $\varphi(P, \succ) \neq TTC(P, \succ)$.

The proof will proceed by induction on the steps of TTC. Let $\tilde{I}_k(P, \succ)$ be the set of agents who are matched in step k of $TTC(P, \succ)$.

Consider agents in $\tilde{I}_1(P, \succ)$. Suppose that $\varphi(P, \succ)(i) \neq TTC(P, \succ)(i)$ for some $i \in \tilde{I}_1(P, \succ)$. Let $c = \{o_k, i_k\}_{k=1,\ldots,K}$ be the cycle in which i is matched with $TTC(P, \succ)(i)$ and $i = i_K$.

Note that every i_k trades o_k for o_{k+1}, which is her first choice. Consider the alternative preference relation $P'_{i_K} : o_1 o_K \ldots.$ By construction, the TTC matching remains the same, i.e.

$$TTC(P'_{i_K}, P_{-\{i_K\}}, \succ) = TTC(P, \succ).$$

Since o_1 is i_K's first choice and

$$\varphi(P, \succ)(i_K) \neq TTC(P, \succ)(i_K),$$

$o_1 P_{i_K} \varphi(P, \succ)(i_K)$. Also, since o_K ranks i_K highest, by recursive individual rationality of φ,

$$\varphi(P'_{i_K}, P_{-\{i_K\}}, \succ)(i_K) R'_{iK} o_K$$

so that

$$\varphi(P'_{i_K}, P_{-\{i_K\}}, \succ)(i_K) \in \{o_1, o_K\}.$$

By strategy-proofness of φ, it must be that

$$\varphi(P'_{i_K}, P_{-\{i_K\}}, \succ)(i_K) = o_K.$$

Then

$$\varphi(P'_{i_K}, P_{-\{i_K\}}, \succ)(i_{K-1}) \neq TTC(P'_{i_K}, P_{-\{i_K\}}, \succ)(i_{K-1}) = o_K.$$

Now consider the alternative preference relation $P'_{i_{K-1}} : o_K o_{K-1} \ldots.$ By construction, the TTC matching remains the same, i.e.

$$TTC(P'_{i_{K-1}}, P'_{i_K}, P_{\{-i_K\}}, \succ) = TTC(P'_{i_K}, P_{-\{i_K\}}, \succ) = TTC(P, \succ).$$

Since o_K is i_{K-1}'s first choice and

$$\varphi(P'_{i_K}, P_{-\{i_K\}}, \succ)(i_{K-1}) \neq TTC(P'_{i_K}, P_{-\{i_K\}}, \succ)(i_{K-1}),$$

we obtain

$$o_K P_{i_{K-1}} \varphi(P'_{i_K}, P_{-\{i_K\}}, \succ)(i_{K-1}).$$

Also, since o_{K-1} ranks i_{K-1} highest, by recursive individual rationality of φ,

$$\varphi(P'_{i_{K-1}}, P'_{i_K}, P_{-\{i_{K-1}, i_K\}}, \succ)(i_{K-1}) R'_{i_{K-1}} o_{K-1}$$

so that

$$\varphi(P'_{i_{K-1}}, P'_{i_K}, P_{-\{i_{K-1}, i_K\}}, \succ)(i_{K-1}) \in \{o_K, o_{K-1}\}.$$

By strategy-proofness of φ, it must be that

$$\varphi(P'_{i_{K-1}}, P'_{i_K}, P_{-\{i_{K-1}, i_K\}}, \succ)(i_K) = o_{K-1}.$$

Then

$$\varphi(P'_{i_{K-1}}, P'_{i_K}, P_{-\{i_{K-1}, i_K\}}, \succ)(i_{K-1}) \neq TTC(P'_{i_{K-1}}, P'_{i_K}, P_{-\{i_{K-1}, i_K\}}, \succ)(i_{K-2}) = o_{K-1}.$$

Repeating this argument recursively for every agent in the cycle c, we obtain that $\varphi(P'_c, P_{-c}, \succ)(i_k) = o_k$ where $P'_c = \{P'_{i_k}\}_{i_k \in c}$ and $P'_{i_k} : o_{k+1} o_k ....$ Then this contradicts with Pereto efficiency of φ because every agent in the cycle will be better off if every i_k is matched with o_{k+1} without changing the matching of agents in $I - c$.

So $\varphi(P, \succ)(i) = TTC(P, \succ)(i)$ for all i that is matched in the first step of TTC, i.e. $i \in \tilde{I}_1(P, \succ)$.

Now assume that $\varphi(P, \succ)(i) = TTC(P, \succ)(i)$ for all i that is matched in the first $m-1$ steps of TTC, i.e. $i \in \bigcup_{l=1,...,m-1} \tilde{I}_l(P, \succ)$. We will show that

$$\varphi(P, \succ)(i) = TTC(P, \succ)(i)$$

for all $i \in \tilde{I}_m(P, \succ)$. Suppose to the contrary that there exists $i \in \tilde{I}_m(P, \succ)$ such that

$$\varphi(P, \succ)(i) \neq TTC(P, \succ)(i).$$

Consider the T_k sets in the definition of recursive individual rationality. By construction, $i \in \tilde{I}_m(P, \succ)$ implies that $i \in T_l$ for some $l = 1, ..., m$. If $i \in T_l$ for some $l = 1, ..., m-1$, then $\varphi(P, \succ)(i) = TTC(P, \succ)(i)$ by our induction hypothesis, a contradiction. So $i \in T_m$. Also, by some overuse of notation, let $c = \{o_k, i_k\}_{k=1,...,K}$ be the cycle in which i is matched with $TTC(P, \succ)(i)$ in step m of $TTC(P, \succ)$ and $i = i_K$. By construction of TTC and the selection i_K, $o_K \in e(i_K)$. So $\varphi(P, \succ)(i_K) R_{i_K} o_K$ by recursive individual rationality. We repeat the arguments above to arrive a contradiction. In the order of $i_K, i_{K-1}, ..., i_1$, replace agents' preference relations one by one with $P'_{i_k} : o_{k+1} o_k ....$ At each replacement, our induction hypothesis, recursive individual rationality and strategy-proofness of φ imply that

$$\varphi(P'_{i_k}, ..., P'_{i_K}, P_{-\{i_k,...,i_K\}}, \succ)(i_k) = o_k$$

and eventually

$$\varphi(P'_c, P_{-c}, \succ)(i_k) = o_k$$

where $P'_c = \{P'_{i_k}\}_{i_k \in c}$, which contradicts with Pareto efficiency of φ.

Therefore $\varphi(P, \succ) = TTC(P, \succ)$ for all $(P, \succ)$.