Robust Exchange under Single-Peaked Preferences*

Yuichiro Kamada† and Yosuke Yasuda‡

March 24, 2025

Abstract

Kamada and Yasuda (2025) consider the standard housing exchange problem of Shapley and Scarf (1974) with a constraint on the cycle size. This paper shows that their impossibility results—there is no mechanism with certain desirable properties—continue to hold under the restricted domain of single-peaked preferences except for some special cases. One important case is where the cycle size is restricted to be at most 2, where we construct a strategy-proof mechanism that always induces a constrained efficient exchange that respects the cycle size constraint.

Keywords: cycle size, efficiency, single-peaked preferences

*Jiarui Xie provided excellent research assistance. All remaining errors are ours.

†University of California, Berkeley, Haas School of Business, Berkeley, 2220 Piedmont Avenue, CA 94720, USA and The University of Tokyo, Faculty of Economics, 7-3-1 Hongo, Bunkyoku, Tokyo, 113-0033, Japan. E-mail: `y.cam.24@gmail.com`

‡Osaka University, Department of Economics, 1-7 Machikaneyama, Toyonaka, Osaka 560-0043 Japan. E-mail: `yosuke.yasuda@gmail.com`

1 Introduction

Kamada and Yasuda (2025) consider the standard housing exchange problem of Shapley and Scarf (1974) with a constraint on the cycle size. Impossibility results hold in such an environment: No mechanism that respects the cycle size constraint and implements all cycles (satisfying the cycle size constraint) assigning every involved agent their first choice is strategy-proof. Also, no mechanism that respects the cycle size constraint and Pareto-efficient (in the set of exchanges satisfying the cycle size constraint) is strategy-proof. The proofs of these results hinge on a cyclic nature of agents' preferences.

In this paper, we consider a restricted domain of single-peaked preferences, which limits the scope of constructing cyclic preferences. Despite the restriction, we prove impossibility results except for some special cases. One important case is where the cycle size is restricted to be at most 2, i.e., only bilateral trades are allowed. In such a case, we construct a strategy-proof mechanism that always induces a constrained efficient exchange that respects the cycle size constraint.

2 Model

Let $I = \{1, 2, \ldots, N\}$ be the set of agents. Each agent i possesses an indivisible object i, and thus I also denotes the set of objects. Each agent i has a strict preference relation $\succ_i$ over objects in I. We write $j \succeq_i j'$ if and only if $j \succ_i j'$ or $j = j'$. We say that object j is **acceptable** to agent i if $j \succeq_i i$. We say that agent i's preferences are **single-peaked** if there is $j^* \in I$ such that $j \succ_i j-1$ for every $j \leq j^*$ and $j \succ_i j+1$ for every $j \geq j^*$. We sometimes let $\succ := (\succ_j)_{j \in I}$ to denote a preference profile.

An **exchange** is a mapping $\mu : I \to I$ that assigns to each agent an object, with a restriction that there is no pair of agents (i, j) such that $\mu(i) = \mu(j)$. For notational simplicity, we denote $\mu(i)$ by μ_i. If $\mu_i = i$, we say that i is **unassigned** under μ. If $\mu_i \neq i$, we say that i **receives an object**.

We say that an exchange μ is **individually rational** if $\mu_i \succeq_i i$ for every $i \in I$. An exchange μ is called **efficient** if there is no other μ' such that $\mu'_i \succeq \mu_i$ for every $i \in I$.

Given an exchange μ, **cycle in** μ is a sequence of agents, $(i_1, i_2, \ldots, i_k)$ for some positive integer k, such that (i) $i_l \neq i_{l'}$ for any pair (l, l') with $l \neq l'$, and (ii) $\mu_{i_l} = i_{l+1}$

for each l where we set $i_{k+1} := i_1$. We say that a cycle in μ is a **k-cycle in μ** if it involves k agents.

Definition 1. An exchange μ is **k-robust** if there is no k'-cycle in μ such that $k' > k$.

We impose constraints on the maximum cycle size: we introduce a threshold k such that we only allow an exchange to be k-robust.

Note that our model coincides with the standard house allocation problem if $k = N$, i.e., exchanges have no size restriction.

Definition 2. Given a preference profile $(\succ_j)_{j \in I}$, a **unanimous trading cycle** is a sequence $(i_1, \ldots, i_k)$ such that (i) $i_{l+1} \succeq_{i_l} i_{l'}$ for every $l' \in \{1, \ldots, N\}$ where we set $i_{k+1} := i_1$, and (ii) $i_l \neq i_{l'}$ for any pair (l, l') with $l \neq l'$.

An exchange μ is **k-unanimous** if, whenever $(i_1, \ldots, i_{k'})$ is a unanimous trading cycle with size $k' \leq k$, we have $\mu_{i_l} = i_{l+1}$ for each l where we set $i_{k'+1} := i_1$.

An exchange μ is called **k-efficient** if it is k-robust and there is no other k-robust μ' such that $\mu'_i \succeq \mu_i$ for every $i \in I$.

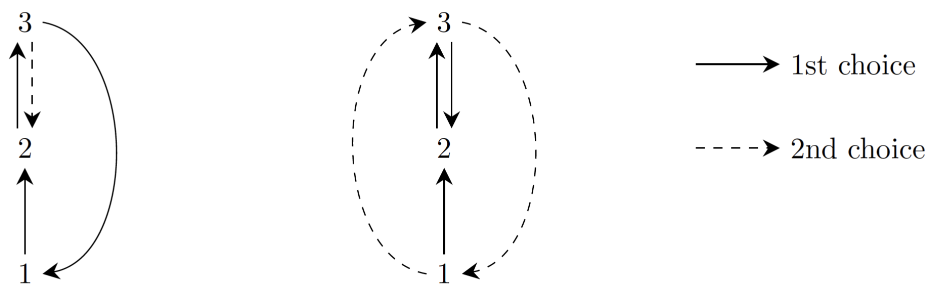
We note that k-unanimity does not imply or is implied by k-efficiency. k-unanimity is not weaker because it only considers the first choices, while k-efficiency also pertains to worse choices. k-efficiency is not weaker because it can give priority to agents who are not anyone's first choice at the expense of others in a unanimous trading cycle with size k. The following example illustrates.

Example 1 (k-unanimity and k-efficiency)**.** Suppose that $I = \{1, 2, 3\}$ and $k = 2$. Consider the following preferences.[1]

$$\succ_1 : 2, 1;$$
$$\succ_2 : 3, 2;$$
$$\succ_3 : 1, 2, 3.$$

Panel (a) of Figure 1 provides a graphical representation of these preferences. Since there is no unanimous trading cycle with size 2 or fewer, the exchange where everyone is unassigned is 2-unanimous. However, such an exchange is not 2-efficient because an alternative exchange where agents 2 and 3 receive each other's object while agent 1 is unassigned is weakly better for everyone.

[1] This is the same example as in Kamada and Yasuda (2025).



(a) Exchange under $\succ$.　　(b) Exchange under $\succ'$.

Figure 1: k-unanimity and k-efficiency. In panel (a), the exchange where everyone is unassigned is 2-unanimous but not 2-efficient. In panel (b), the exchange where agents 1 and 3 receive each other's object is not 2-unanimous but 2-efficient.

Now, consider the following preferences.

$$\succ'_1 : 2, 3, 1;$$
$$\succ'_2 : 3, 2;$$
$$\succ'_3 : 2, 1, 3.$$

Panel (b) of Figure 1 provides a graphical representation of these preferences. The exchange where agents 1 and 3 receive each other's object while agent 2 is unassigned is 2-efficient because any other exchange would make either of agents 1 and 2 unassigned, which would make them worse off. However, such an exchange is not 2-unanimous because agents 2 and 3 constitute a unanimous trading cycle with size 2. □

A mechanism is a function that assigns to each preference profile an exchange. We denote the object that agent i receives by $\psi_i((\succ_j)_{j\in I})$. A mechanism is called **individually rational** and **efficient**, respectively, if it always returns an individually rational exchange and an efficient exchange. We say that a mechanism ψ is **strategy-proof under single-peaked preferences** if, for every single-peaked preference profile $(\succ_j)_{j\in I}$ and every agent $i \in I$,

$$\psi_i((\succ_j)_{j\in I}) \succeq_i \psi_i(\succ'_i, (\succ_j)_{j\neq i})$$

holds for every $\succ'_i$. Note that this definition requires that the mechanism is immune to misreporting of *any* preferences, including those that are not single-peaked. This means that the conclusions of the negative results we will state (Theorems 1 and 2) are weaker under this definition than under an alternative definition that would only require immunity against misreported preferences that are single-peaked.[2] On the other hand, that of the positive result (Theorem 3) is stronger under our definition.

Definition 3. A mechanism is

1. k**-robust** if it returns a k-robust exchange given any input;
2. k**-unanimous** if it returns a k-unanimous exchange given any input;
3. k**-efficient** if it returns a k-efficient exchange given any input.
4. k**-efficient under single-peaked preferences** if it returns a k-efficient exchange given any single-peaked preferences.

3 Results

Theorem 1. *Fix any N and k such that $1 < k < N$. Then, no k-robust mechanism is k-unanimous and strategy-proof under single-peaked preferences if and only if $N > 3$.*

To show the "if" direction, we construct examples. We leave the details on this in the Appendix. To show the "only if" direction, we consider the case of $k = 2$ and $N = 3$. In this case, the following is a 2-robust mechanism that is 2-unanimous and strategy-proof under single-peaked preferences:

Count the number of pairs such that in each pair, the involved agents regard each other's object as acceptable. There are the following four cases.

1. 0 pairs: Every agent is unassigned.
2. 1 pair: The agents in the pair receive the object of the other agent in the pair, while the remaining one agent is unassigned.
3. 2 pairs: Then there must be a unique agent i who is involved in those two pairs, and we have such an agent receive her first choice object j and j receives object i, while the remaining agent is unassigned.

[2]Although we view this definition is a plausible one, we also note that the conclusion of Theorem 1 holds even under the alternative definition.

4. 3 pairs (i.e., each agent regards the other two objects as acceptable): There are either zero or one 2-unanimous cycle.

   (a) If there exist zero 2-unanimous cycles, then agent 3 receives her first choice j, and agent j receives objet 3.[3] The remaining one agent is unassigned.

   (b) If there exists one 2-unanimous cycle, then implement that cycle, and the remaining one agent is unassigned.

It is straightforward to see, by inspection, that this mechanism is 2-robust and 2-unanimous. The proof in the Appendix shows that the mechanism is also strategy-proof under single-peaked preferences.

Theorem 2. *No k-robust mechanism is individually rational, k-efficient, and strategy-proof under single-peaked preferences if $k > 2$.*

To show the "if" direction, we construct examples. We leave the details on this in the Appendix.

We do not know if the converse holds, but we show that a weakening of the converse holds where we replace 2-efficiency with 2-efficiency under single-peaked preferences. To see this, we construct a 2-robust mechanism that is individually rational, 2-efficient, and strategy-proof under single-peaked preferences. The mechanism is based on the following Up-Down algorithm:

Up-Down Algorithm:

Step 0: Given the preference profile $\succ$, partition I into three sets:

$$U := \{j \in I | \succ_j \text{ is single-peaked and } \exists j' > j \text{ s.t. } j' \succ_j j\};$$

$$D := \{j \in I | \succ_j \text{ is single-peaked and } \exists j' < j \text{ s.t. } j' \succ_j j\};$$

$$M := I \setminus (U \cup D).$$

Note that, if $\succ_j$ is single-peaked, then j cannot simultaneously belong to U and D.

Step 1: Consider the highest agent in U, and denote her by i^1, i.e., we have $i^1 \geq j$ for every $j \in U$. Consider the set $S^1 := \{j \in D | i_1 \succ_j j\}$. Define $T^1 := \{j \in D | i^1 < j\} \setminus S^1$. Let $\mu_j = j$ for all $j \in T^1$.

[3] It can be agent 1 who gets priority here, but it cannot be agent 2. This point will become clear in the proof in the Appendix (see footnote 5).

- If S^1 is empty, let $\mu_{i^1} = i^1$. Let $V^1 = \{i^1\} \cup T^1$. Go to Step 2.
- If S^1 is nonempty, find the best object for i^1 in S^1, and denote it by $c(i^1)$.
  - If $i^1 \succ_{i^1} c(i^1)$, then let $\mu_{i^1} = i^1$. Let $V^1 = \{i^1\} \cup T^1$. Go to Step 2.
  - If $c(i^1) \succ_{i^1} i^1$, then let $\mu_{i^1} = c(i^1)$ and $\mu_{c(i^1)} = i^1$. Let $V^1 = \{i^1, c(i^1)\} \cup T^1$. Go to Step 2.

Step l: If $l = |U| + 1$, let $\mu_j = j$ for any agent j for whom the algorithm has not specified the object to be assigned, and terminate the algorithm and output μ.

Otherwise, consider the l-th highest agent in U, and denote her by i^l. Consider the set $S^l := \{j \in D | i_l \succ_j j\} \setminus V^{l-1}$. Define $T^l := \{j \in D | i^l < j \text{ and } j \succ_j i^l\}$.

- If S^l is empty, let $\mu_{i^l} = i^l$. Let $V^l = V^{l-1} \cup \{i^l\} \cup T^l$. Go to Step $l + 1$.
- If S^l is nonempty, find the best agent for i^l in S^l, and denote him by $c(i^l)$.
  - If $i^l \succ_{i^l} c(i^l)$, then let $\mu_{i^l} = i^l$. Let $V^l = V^{l-1} \cup \{i^l\} \cup T^l$. Go to Step $l + 1$.
  - If $c(i^l) \succ_{i^l} i^l$, then let $\mu_{i^l} = c(i^l)$ and $\mu_{c(i^l)} = i^l$. Let $V^l = V^{l-1} \cup \{i^l, c(i^l)\} \cup T^l$. Go to Step $l + 1$.

The following theorem is used to show the "only if" direction of Theorem 2.

Theorem 3. *The up-down algorithm is* 2*-robust, individually rational,* 2*-efficient under single-peaked preferences, and strategy-proof under single-peaked preferences.*

The idea of the algorithm is that the single-peakedness implies that each agent i's acceptable objects are such that either (i) all of them are no smaller than i or (ii) all of them are no greater than i. Given this fact, we consider a type of serial dictatorship where the highest agent among those whose acceptable objects are above themselves becomes the first dictator. If, for example, j is this dictator's first choice but she does not deem the dictator acceptable, then j is determined to be unassigned, i.e., she receives object j. This construction works because the single-peakedness implies that, if j deems the dictator unacceptable, then j, who we know is above the dictator because the dictator is in U, will deem all the future dictators (who will be below the first dictator) unacceptable.

It may be instructive to consider the relationship with a two-sided matching model. (Ignoring the agents who regard only herself as acceptable) One can imagine that, according to the reported preferences, the agents in (i) are put on one side and the agents in (ii) are put on the other side. An agent on one side can only be assigned an object of the agent on the other side. Although which side an agent belongs to would depend on her reported preferences, it is straightforward to see that no agent has an incentive to be on the "wrong side." The Up-Down mechanism corresponds to the standard serial dictatorship mechanism where the dictators are always on the same side. Such a mechanism is not strategy-proof in the standard two-sided matching model because the agents on the other side may have an incentive to report that an early dictator is unacceptable in order to be matched with a later dictator. In our model, however, this type of misreporting will not be profitable due to the single-peaked nature of preferences and the order of dictators: If an agent named by an early dictator i (who must be above i because $i \in U$) reports that i would be unacceptable, then the same agent will be treated as regarding all later dictators (who must be below i) as unacceptable.[4]

A Proof of Theorem 1

The "only if" direction:

Consider the mechanism described after the statement of Theorem 1.

Consider agent $i \in \{1, 2, 3\}$. Suppose first that agent i regards no object but object i as acceptable. Notice that, if she reports that object i is the uniquely acceptable object, then she receives object i under the above mechanism. Hence, truthful preference report is optimal for i in this case.

Second, suppose that agent i regards object $j \neq i$ as her first choice. Let $k \in \{1, 2, 3\} \setminus \{i, j\}$. Then, one of the following happens.

1. If object i is agent j's first choice, then if agent i reports that object j is her first choice, then agent i receives object j under the above mechanism. Hence, truthful preference report is optimal for i in this case.

[4] A similar analogy could be drawn for the case of $k = 2$ in Theorem 1. In the context of two-sided matching, Takagi and Serizawa (2010) show that no strategy-proof mechanism is individually rational and matches all the mutual first choices (Takagi and Serizawa (2010) call this property "2-unanimity"). This result is shown when there are at least two agents *on each side*, and accordingly, we have the impossibility when $N \geq 4$. In contrast to the case of Theorem 2, however, single-peakedness of preferences does not help to eliminate the incentive of profitable deviation in this case.

2. If object i is agent j's second choice, then if agent i reports that object j is her first choice, the only case in which agent i does not receive object j is when agent k regards object j as acceptable. In such a case, either (i) all three agents regard each other's object as acceptable and such a preference report profile (i.e., all agents regard each other's object as acceptable) arises only when some agent other than i reports a non-single-peaked preferences, or (ii) agents j and k receive each other's object and there is no preference report of agent i that makes her receive object j.[5] Hence, truthful preference report is optimal for i in this case.

3. If j reports that object i is not acceptable, then i will not be assigned object j irrespective of her preference report. Moreover, whether agent i is assigned object k depends only on agents j and k's preference report and whether agent i reports that object k as acceptable or not. Hence, truthful preference report is optimal for i in this case.

This completes the proof for the "only if" direction.

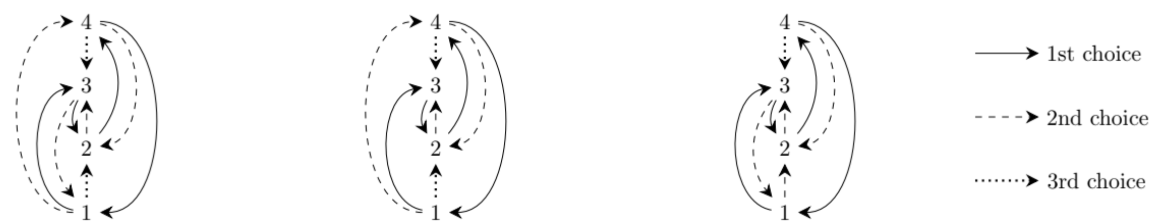
The "if" direction:

We first prove the "if" direction of Theorem 1 for the simplest cases of $k = 2$ and $k = 3$ (and $N \geq 4$) to illustrate the intuition. Generalizing this proof, we then provide the proof for the fully general case.

Proof of the "if" direction of Theorem 1 for $k = 2$ and $k = 3$ (and $N \geq 4$). Suppose that $k = 2$ or $k = 3$ (and $N \geq 4$). The following proof works for either case. Consider a subset of agents, denoted by $S \subseteq I$, which consists of 4 agents: $S = \{1, 2, 3, 4\}$. Such a subset can be taken because $4 \leq N$.

Consider the following three different preference profiles, $\succ$, $\succ^A$, and $\succ^B$. Figure 2 provides a graphical representation of these preferences. Note that everyone has single-peaked preferences under any of these preference profiles.

[5] If $i = 3$ and i regards object k as unacceptable, one might think that by reporting that object k is acceptable, i can create a preference profile with 3 pairs, thereby giving agent i the right to become a dictator to receive object j. However, this implies that either $j = 2$ or $k = 2$, and their preference reports of regarding three objects as acceptable violates single-peakedness. Since strategy-proofness under single-peaked preferences does not require agent i to have an incentive for truthful reporting when some other agents report non-single-peaked preferences, one can ignore this case.



(a) Exchange under $\succ$. (b) Exchange under $\succ^A$. (c) Exchange under $\succ^B$.

Figure 2: The counterexample for the case of $k = 2$ or $k = 3$ in the proof of Theorem 1.

.

Preference profile $\succ = (\succ_1, \succ_2, \succ_3, \succ_4)$, where

$$\succ_1 : 3, 4, 2, 1$$
$$\succ_2 : 4, 3, 2$$
$$\succ_3 : 2, 1, 3$$
$$\succ_4 : 1, 2, 3, 4$$

Preference profile $\succ^A = (\succ_1, \succ_2, \succ_3^A, \succ_4)$, where

$$\succ_3^A : 2, 3.$$

Preference profile $\succ^B = (\succ_1^B, \succ_2, \succ_3, \succ_4)$, where

$$\succ_1^B : 3, 2, 1.$$

Let ψ be a 3-robust mechanism that is 3-unanimous and strategy-proof. First, note that under any of $\psi(\succ)$, $\psi(\succ^A)$, and $\psi(\succ^B)$, if a given agent i has two acceptable objects besides object i and the second-choice object j is such that agent j regards object i as his first choice, then agent i must receive her first- or second-choice object. This is because, if agent i does not receive such an object, she could misreport to say that her second choice was her first choice, thereby creating a unanimous trading cycle of size 2. This implies that all agents receive an object under $\psi(\succ)$, agents 1, 2, and 4 receive an object under $\psi(\succ^A)$, and agents 2, 3, and 4 receive an object under $\psi(\succ^B)$.

Step 1: Under $\psi(\succ^A)$, it is not possible to have agents 1, 2, and 4 receive an object while agent 3 is unassigned because that would mean that agent 1 receives object 2, which is her third choice. This implies that all agents receive an object. Since object 2 is the only acceptable object for agent 3, $\psi(\succ^A) =: \mu^A$ must be such that $(\mu_1^A, \mu_2^A, \mu_3^A, \mu_4^A) = (4, 3, 2, 1)$.

Step 2: Under $\psi(\succ^B)$, it is not possible to have agents 2, 3, and 4 receive an object while agent 1 is unassigned. This implies that all agents receive an object. Since object 4 is not acceptable by agents 1 or 3, $\psi(\succ^B) =: \mu^B$ must be such that $(\mu_1^B, \mu_2^B, \mu_3^B, \mu_4^B) = (3, 4, 1, 2)$.

Step 3: We show that an agent with an incentive to misreport her preference must exist under $\succ$. Since agent 2 does not regard object 1 as acceptable and agent 3 does not regard object 4 as acceptable, $\psi(\succ)$ must be equal to either μ^A or μ^B.

If $\psi(\succ) = \mu^B$, then agent 3 becomes better off by misreporting her preferences to say $\succ_3^A$. If $\psi(\succ) = \mu^A$, then agent 1 becomes better off by misreporting her preferences to say $\succ_1^B$.

This completes the proof for the "if" direction for the cases of $k = 2$ and $k = 3$ (and $N \geq 4$). □

Proof of the "if" direction of Theorem 1 for $k \geq 4$. Suppose that $k \geq 4$. Consider a subset of agents, denoted by $S \subseteq I$, which consists of $k+1$ agents: $S = \{1, 2, 3_1, 3_2, \ldots, 3_m, 4\}$, where $m = k - 2$. Such a subset can be taken because $k < N$.

Consider the following three different preference profiles, $\succ$, $\succ^A$, and $\succ^B$. Figure 3 provides a graphical representation of these preferences. Note that everyone has single-peaked preferences under any of these preference profiles.

Preference profile $\succ = (\succ_1, \succ_2, \succ_{3_1}, \ldots, \succ_{3_m}, \succ_4)$, where

$$\begin{aligned}
\succ_1 &: 3_m, 4, 3_{m-1}, \ldots, 3_1, 2, 1 \\
\succ_2 &: 4, 3_m, \ldots, 3_1, 2 \\
\succ_{3_1} &: 2, 3_1 \\
\succ_{3_l} &: 3_{l-1}, 3_l \text{ for all } l \in \{2, \ldots, m-1\} \\
\succ_{3_m} &: 3_{m-1}, 3_{m-2}, 3_m \\
\succ_4 &: 1, 2, 3_1, \ldots, 3_m, 4
\end{aligned}$$

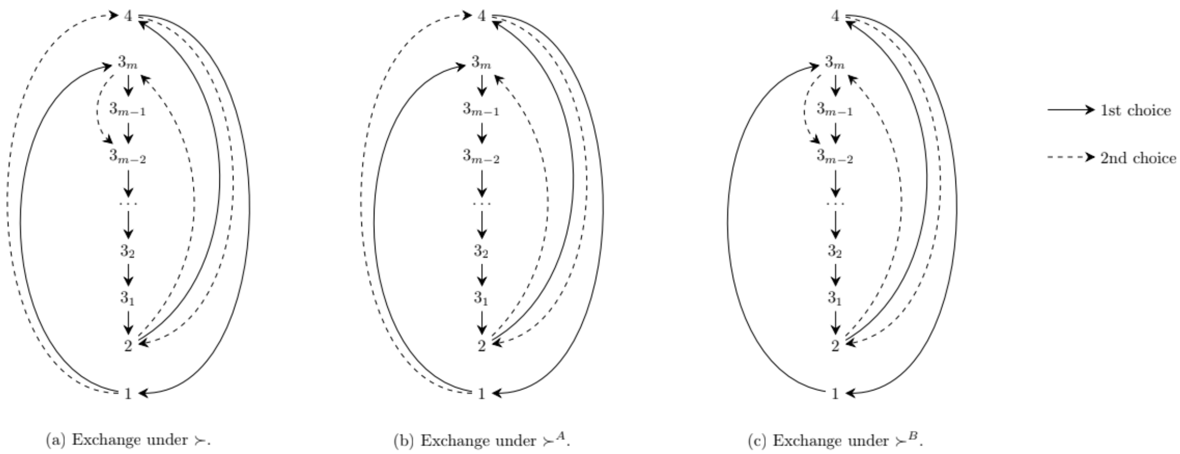


Figure 3: The counterexample for the case of $k \geq 4$ in the proof of Theorem 1.

where 3_0 is defined to be 2.

Preference profile $\succ^A = (\succ_1, \succ_2, \succ_{3_1}, \ldots, \succ^A_{3_m}, \succ_4)$, where

$$\succ^A_{3_m}: 3_{m-1}, 3_m.$$

Preference profile $\succ^B = (\succ^B_1, \succ_2, \succ_{3_1}, \ldots, \succ_{3_m}, \succ_4)$, where

$$\succ^B_1: 3_m, 3_{m-1}, \ldots, 3_1, 2, 1.$$

Let ψ be a k-robust mechanism that is k-unanimous and strategy-proof, where $k = m + 2$.

Consider any preference profile in $\{\succ, \succ^A, \succ^B\}$.

We first make the following claim.

Claim 1. *Under any preference profile in $\{\succ, \succ^A, \succ^B\}$, if a given agent has multiple acceptable objects besides object i, then she must receive an object that is her first or second choice under ψ.*

This holds because otherwise, they could claim their second choice to be the first choice, by which a unanimous trading cycle of size no greater than $m + 2$ exists.

Consider any preference profile in $\{\succ, \succ^A\}$. $l \in \{l' \in \mathbb{N} | 2 \leq l' \leq m\} \setminus \{m - 1\}$, and suppose that agent 3_l is unassigned. We show that 3_{l-1} is unassigned, too. To see this, note that 3_{l-1} is acceptable only by agents 1, 2, 3_{l-1}, 3_l, and 4.

For agents 1, 2, and 4, object 3_{l-1} is worse than their second choice. Hence, by Claim 1 and the assumption that agent 3_l is unassigned, object 3_{l-1} is not received by anyone. Hence, 3_{l-1} is unassigned.

Now we show that agent 3_{m-2} is unassigned if both 3_m and 3_{m-1} are unassigned.

1. Consider the preference profile $\succ^A$. Suppose that 3_{m-1} is unassigned. We show that 3_{m-2} is unassigned as well. To see this, note that object 3_{m-2} is acceptable only by agents 1, 2, 3_{m-2}, 3_{m-1}, and 4. First, suppose that $m > 2$. Then, for agents 1, 2, and 4, object 3_{m-2} is worse than their second choice. Hence, by Claim 1 and the assumption that agent 3_{m-1} is unassigned, object 3_{m-2} is not received by anyone. Hence, agent 3_{m-2} is unassigned.

   If $m = 2$, for agent 1, object 3_{m-2} $(= 2)$ is worse than her second choice. Hence, the only possibility is for agent 4 to receive object 2. This implies that object 1 is not received by any agent. Hence, agent 1 is unassigned. This contradicts Claim 1.

2. Consider the preference profile $\succ$. Suppose that both agents 3_m and 3_{m-1} are unassigned. We show that agent 3_{m-2} is unassigned as well. To see this, note that object 3_{m-2} is acceptable only by agents 1, 2, 3_{m-2}, 3_{m-1}, 3_m, and 4. First, suppose that $m > 2$. Then, for agents 1, 2, and 4, object 3_{m-2} is worse than their second choice. Hence, by Claim 1 and the assumption that both agents 3_m and 3_{m-1} are unassigned, object 3_{m-2} is not received by anyone. Hence, agent 3_{m-2} is unassigned.

   If $m = 2$, for agent 1, object 3_{m-2} $(= 2)$ is worse than her second choice. Hence, the only possibility is for agent 4 to receive object 2. This implies that object 1 is not received by any agent. Hence, agent 1 is unassigned. This contradicts Claim 1.

Overall, the above arguments and the fact that every agent i in $\{3_2, \ldots, 3_{m-1}\}$ has only a single acceptable object besides object i imply that, for any preferences in $\{\succ, \succ^A\}$, one of the following holds.

(i). All agents in $\{3_1, \ldots, 3_m\}$ are unassigned.

(ii). All agents in $\{3_1, \ldots, 3_m\}$ receive an object.

(iii). Agent 3_{m-1} is unassigned, while all agents in $\{3_1, \ldots, 3_{m-2}, 3_m\}$ receive an object.

Consider case (i). Consider any preference profile in $\{\succ, \succ^A\}$. By Claim 1, both agents 1 and 2 must receive object 4, but that is infeasible. Hence, case (i) cannot happen under these preferences.

This implies that the only remaining cases are (ii) and (iii) for each of preference profile in $\{\succ, \succ^A\}$. For each preference profile in $\{\succ, \succ^A\}$, agent 3_1 receives an object in (ii) and (iii), and object 2 is the only acceptable choice for agent 3_1. These facts imply that agent 3_1 receives 2.[6]

Now, consider the preference profile $\succ$. Note that Claim 1 implies that agent 1 receives either object 3_m or 4 under $\succ$.

1. Suppose that agent 1 receives object 3_m. Then, Claim 1 implies that agent 2 receives object 4. Thus, under case (ii), object 3_{m-1} must be received by 3_m. But this implies that $\psi(\succ)$ is not $(m+2)$-robust. Hence, case (iii) happens. This means that agent 3_m receives object 3_{m-2}. In this case, suppose that agent 3_m misreports to say $\succ^A_{3_m}$. Then, since we are in case (ii) or case (iii) under $\succ^A$, 3_m receives object 3_{m-1}, and hence such a misreporting strictly improves agent 3_m's assignment.[7] This contradicts the assumption that ψ is strategy-proof.

2. Suppose that agent 1 receives object 4. Suppose that agent 1 misreports to say $\succ^B_1$. Then, if 1 receives object 3_{m-1}, agent 3_{m-1} must receive some object. Since object 3_{m-2} is the only acceptable object for agent 3_{m-1}, agent 3_{m-1} receives object 3_{m-2}. But this implies that 3_m is unassigned, which contradicts Claim 1. Again by Claim 1, agent 1 must receive either object 3_m or 3_{m-1}, so this implies that agent 1 receives object 3_m. Hence, agent 1's misreporting of $\succ^B_1$ strictly improves agent 1's assignment. This contradicts the assumption that ψ is strategy-proof. Contradiction.

Overall, neither cases (ii) nor (iii) can happen under $\succ$. Therefore, there is no exchange μ such that $\psi(\succ) = \mu$. Contradiction.

This completes the proof for the "if" direction for the cases of $k \geq 4$. $\square$

[6]The argument needs a minor modification when $m = 2$, but the result goes through regardless.

[7]More specifically, case (iii) would not happen under $\succ^A$ because 3_m receiving an object implies that 3_m receives object 3_{m-1}, which contradicts agent 3_{m-1} being unassigned under case (iii).

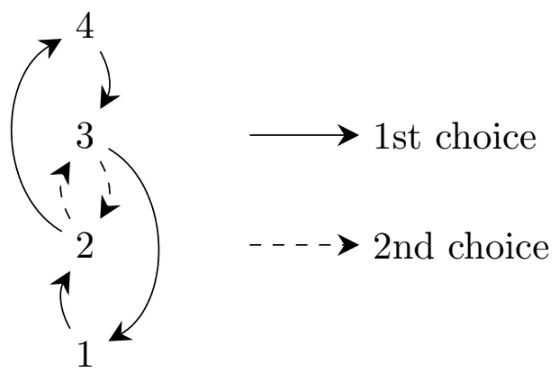


Figure 4: The counterexample for the case of $k = 3$ in the proof of Theorem 2.

A.1 Proof of Theorem 2

Note that the "only if" direction of Theorem 2 is an implication of Theorem 3.

For the "if" direction, we first prove it for the simplest cases of $k = 3$ to illustrate the intuition. Generalizing this proof, we then provide the proof for the fully general case.

Proof of the "if" direction of Theorem 2 for $k = 3$. Suppose that $k = 3$. Consider a subset of agents, denoted by $S \subseteq I$, which consists of 4 agents: $S = \{1, 2, 3, 4\}$. Such a subset can be taken because $3 = k < |I|$.

Consider the following three different preference profiles, $\succ$, $\succ^A$, and $\succ^B$. Figure 4 provides a graphical representation of these preferences. Note that everyone has single-peaked preferences under $\succ$.

Preference profile $\succ = (\succ_1, \succ_2, \succ_3, \succ_4)$, where

$$\succ_1 : 2, 1$$

$$\succ_2 : 4, 3, 2$$

$$\succ_3 : 1, 2, 3$$

$$\succ_4 : 3, 4$$

Preference profile $\succ^A = (\succ_1, \succ_2, \succ_3^A, \succ_4)$, where

$$\succ_3^A : 1, 3.$$

Preference profile $\succ^B = (\succ_1, \succ_2^B, \succ_3, \succ_4)$, where

$$\succ_2^B: 4, 2.$$

Suppose that ψ is a 3-robust mechanism that is individually rational, 3-efficient, and strategy-proof under single-peaked preferences.

Step 1: Define μ^A by $(\mu_1^A, \mu_2^A, \mu_3^A, \mu_4^A) = (2, 3, 1, 4)$. Under $\succ^A$, μ^A is the only 3-robust exchange that satisfies individual rationality and 3-efficiency. Thus, we must have $\psi(\succ^A) = \mu^A$.

Step 2: Define μ^B by $(\mu_1^B, \mu_2^B, \mu_3^B, \mu_4^B) = (1, 4, 2, 3)$. Under $\succ^B$, μ^B is the only 3-robust exchange that satisfies individual rationality and 3-efficiency. Thus, we must have $\psi(\succ^B) = \mu^B$.

Step 3: We show that an agent with an incentive to misreport her preferences must exist under $\succ$. Under $\succ$, μ^A and μ^B are the only 3-robust exchanges that satisfy individual rationality and 3-efficiency, and thus we must have $\psi(\succ) = \mu^A$ or $\psi(\succ) = \mu^B$.

If $\psi(\succ) = \mu^A$, then agent 2 becomes better off by misreporting his preferences to say $\succ_2^B$.

If $\psi(\succ) = \mu^B$, then agent 3 becomes better off by misreporting her preferences to say $\succ_3^A$.

This completes the proof for the case of $k = 3$. □

Proof of the "if" direction of Theorem 2 for $k \geq 4$. Suppose that $k \geq 4$. Consider a subset of agents, denoted by $S \subseteq I$, which consists of $k+1$ agents: $S = \{1, 2, 3, 3_1, 3_2, \ldots, 3_m, 4\}$, where $m = k - 3$. Such a subset can be taken because $k < N$.

Consider the following three different preference profiles, $\succ$, $\succ^A$, and $\succ^B$. Figure 5 provides a graphical representation of these preferences. Note that everyone has single-peaked preferences under $\succ$.

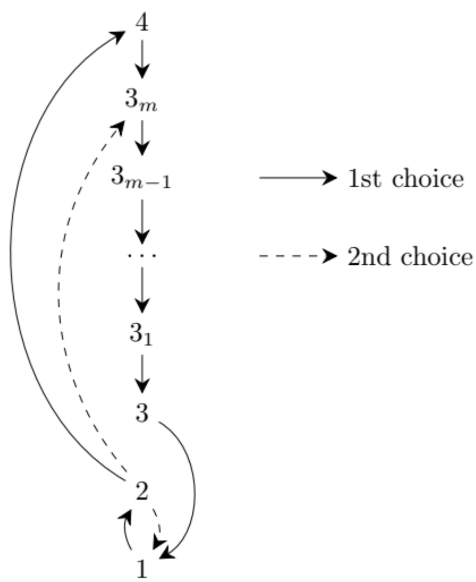


Figure 5: The counterexample for the case of $k \geq 4$ in the proof of Theorem 2.

Preference profile $\succ = (\succ_1, \succ_2, \succ_3, \succ_{3_1}, \ldots, \succ_{3_m}, \succ_4)$, where

$$
\begin{aligned}
&\succ_1 : 2, 1 \\
&\succ_2 : 4, 3_m, \ldots, 3_1, 3, 2 \\
&\succ_3 : 1, 2, 3 \\
&\succ_{3_1} : 3, 3_1 \\
&\succ_{3_l} : 3_{l-1}, 3_l \quad \text{for all } l \in \{2, \ldots, m\} \\
&\succ_4 : 3_m, 4
\end{aligned}
$$

Preference profile $\succ^A = (\succ_1, \succ_2, \succ_3^A, \succ_{3_1}, \ldots, \succ_{3_m}, \succ_4)$, where

$$\succ_3^A : 1, 3.$$

Preference profile $\succ^B = (\succ_1, \succ_2^B, \succ_3, \succ_{3_1}, \ldots, \succ_{3_m}, \succ_4)$, where

$$\succ_2^B : 4, 2.$$

Suppose that ψ is a k-robust mechanism that is individually rational, k-efficient, and strategy-proof under single-peaked preferences, where $k = m + 3$.

Step 1: Define μ^A by

$$\begin{pmatrix} \mu_1^A & \mu_2^A & \mu_3^A & \mu_{3_1}^A & \mu_{3_2}^A & \dots & \mu_{3_m}^A & \mu_4^A \\ 2 & 3_m & 1 & 3 & 3_1 & \dots & 3_{m-1} & 4 \end{pmatrix}$$

by which we mean that for each l, the $(1, l)$ entry is equal to the $(2, l)$ entry. That is, for example, agent 1 receives object 2 under exchange μ^A.

Under $\succ^A$, μ^A is the only $(m+3)$-robust exchange that satisfies individual rationality and $(m+3)$-efficiency. Thus, we must have $\psi(\succ^A) = \mu^A$.

Step 2: Define μ^B by

$$\begin{pmatrix} \mu_1^B & \mu_2^B & \mu_3^B & \mu_{3_1}^B & \mu_{3_2}^B & \dots & \mu_{3_m}^B & \mu_4^B \\ 1 & 4 & 2 & 3 & 3_1 & \dots & 3_{m-1} & 3_m \end{pmatrix},$$

where again, we mean that for each l, the $(1, l)$ entry is equal to the $(2, l)$ entry. Under $\succ^B$, μ^B is the only $(m+3)$-robust exchange that satisfies individual rationality and $(m+3)$-efficiency. Thus, we must have $\psi(\succ^B) = \mu^B$.

Step 3: We show that an agent with an incentive to misreport her preferences must exist under $\succ$. Under $\succ$, μ^A and μ^B are the only $(m+3)$-robust exchanges that satisfy individual rationality and $(m+3)$-efficiency, and thus we must have $\psi(\succ) = \mu^A$ or $\psi(\succ) = \mu^B$.

If $\psi(\succ) = \mu^A$, then agent 2 becomes better off by misreporting his preferences to say $\succ_2^B$.

If $\psi(\succ) = \mu^B$, then agent 3 becomes better off by misreporting her preferences to say $\succ_3^A$.

This completes the proof for the case of $k \geq 4$. □

References

Kamada, Y. and Y. Yasuda (2025): "Robust Exchange," Mimeo.

Shapley, L. and H. Scarf (1974): "On cores and indivisibility," *Journal of mathematical economics*, 1, 23–37.

Takagi, S. and S. Serizawa (2010): "An impossibility theorem for matching problems," *Social Choice and Welfare*, 35, 245–266.