

# Strategyproof Exchange with Multiple Private Endowments

Taiki Todo, Haixin Sun, Makoto Yokoo

Dept. of Informatics

Kyushu University

Motoooka 744, Fukuoka, JAPAN

{*todo, sunhaixin, yokoo*}@agent.inf.kyushu-u.ac.jp

## Abstract

We study a mechanism design problem for exchange economies where each agent is initially endowed with a set of indivisible goods and side payments are not allowed. We assume each agent can withhold some endowments, as well as misreport her preference. Under this assumption, strategyproofness requires that for each agent, reporting her true preference with revealing all her endowments is a dominant strategy, and thus implies individual rationality. Our objective in this paper is to analyze the effect of such private ownership in exchange economies with multiple endowments. As fundamental results, we first show that the revelation principle holds under a natural assumption and that strategyproofness and Pareto efficiency are incompatible even under the lexicographic preference domain. We then propose a class of exchange rules, each of which has a corresponding directed graph to prescribe possible trades, and provide necessary and sufficient conditions on the graph structure so that they satisfy strategyproofness.

## Introduction

The housing market problem, introduced by Shapley and Scarf (1974), is a fundamental exchange model where each good is indivisible and monetary transfers are not allowed. Precisely, there is a group of agents, each of whom is initially endowed with one good, say a house, and has a strict ordering, so-called a *preference*, over the set of all goods owned by the agents. An exchange rule (aka., a *mechanism*) takes the preferences revealed by the agents as input and determines which goods are traded. The model has been applied to various market environments, such as kidney exchanges (Roth, Sönmez, and Ünver 2004) and on-campus university housing markets (Chen and Sönmez 2002).

One desirable property for a mechanism is *strategyproofness*, which requires that for every agent, the truth revelation of her private information, say a preference over a set of houses, to the mechanism is a dominant strategy. From the revelation principle, it is without loss of generality to focus on strategyproof direct revelation mechanisms if we are only interested in exchange rules with dominant strategy equilibria. Fortunately, for the housing market problem, *Gale's Top-Trading-Cycle* (TTC) rule is strategyproof, as well as Pareto

efficient and individually rational. Moreover, it is the only rule that satisfies those three properties (Ma 1994). Due to these advantages, TTC has been attracting much attention from both economists and computer scientists.

In this paper we consider exchange economies where each agent is endowed with a set of multiple goods, instead of a single good. Under that model, Pareto efficiency, individual rationality, and strategyproofness cannot be simultaneously satisfied (Sönmez 1999). Therefore, one main research direction on exchanges with multiple endowments is to achieve strategyproof rules that guarantee a limited notion of efficiency. In particular, Pápai (2003) defined a class of exchange rules and gave a characterization by strategyproofness and individual rationality with some other weak efficiency requirements.

Another important assumption in this work is that each agent's endowments are also private information, and agents reveal to an exchange rule their endowments, as well as their preferences. Since an exchange rule cannot exactly observe which agent really owns which goods, guaranteeing strategyproofness seems much harder than in a traditional case where ownerships are recognized. Indeed, in environments with private (and multiple) goods, the possibility of various manipulations *via endowments* has been pointed out (Postlewaite 1979; Atlamaz and Klaus 2007), such as *hiding* (or withholding). Furthermore, in some anonymous environments, *splitting* an account into multiple ones might also be problematic (Moulin 2008).

In this paper we first focus on *hiding manipulations* as well as misreporting preferences, since they can be easily done by only one agent without any side communication with other agents. We say an exchange rule is strategyproof if for every agent, reporting her true preference with revealing all her endowments (i.e., not hiding anything) to the rule is a dominant strategy. The main objective of this paper is to analyze the effect of such private ownership in exchange economies with multiple endowments. More precisely, we clarify how the space of possible exchanges by strategyproof rules shrinks due to the lack of information on ownership. One of the most closely related works to this paper is Atlamaz and Klaus (2007), which also investigates the effect of hiding manipulations. However, they only focus on hidings and ignore any misreports of preferences.

**Our Results** We investigate exchange problems with multiple private endowments. In our model an exchange rule is *strategyproof* if, for each agent, reporting her true preference with revealing all her endowments is a dominant strategy. The definition implies individual rationality, since hiding all endowments is equivalent to not participating. We prove that the revelation principle holds under a natural assumption, which is regarded as a variation of Yu’s result (Yu 2011). We also show that strategyproofness and Pareto efficiency are incompatible, even if we focus on a “smallest” preference domain called the *lexicographic* preference domain.

We next introduce a class of exchange rules called the *agreement cycles rules*. They are motivated from *segmented trading cycles rules* (Pápai 2003), so that each of them is defined based on a corresponding directed graph called a *one-for-one trading possibility graph*. The performances of these rules are fairly better than trivial strategyproof rules. Under each preference domain, we provide a necessary and sufficient condition on the structure of a one-for-one trading possibility graph for the exchange rule to be strategyproof. We also show that these rules are *split-proof* if and only if they are strategyproof, where split-proofness requires that for each agent, using only one identity is a dominant strategy, even if it can use multiple identities and join an exchange rule multiple times under them.

**Related Works** The Shapley-Scarf housing market (Shapley and Scarf 1974) has various applications, such as roommate problems and kidney exchanges. Ma (1994) characterized TTC for the problem by strategyproofness, Pareto efficiency, and individual rationality. Roth and Postlewaite (1977) showed that it always chooses the unique core assignment. A quite similar model is assignment problems (Pápai 2000), where there are no initial endowments and monetary transfers are still not allowed.

A natural extension of the Shapley-Scarf model is one where each agent is allowed to initially have multiple goods. When at least one agent initially has more than one good, no rule is strategyproof, Pareto efficient, and individually rational simultaneously (Sönmez 1999). Pápai (2007) introduced *fixed deal exchange rules*, which also closely resemble our proposed rules. The main difference is that, while in her model one agent can join more than one deal, that is not the case in ours. Since in our model an exchange rule cannot observe the ownership information, making the whole market a single fixed deal seems the only “consistent” way to guarantee strategyproofness.

In mechanism design literature, variants of strategic manipulations have been studied. Atlamaz and Klaus (2007) studied hidings in an exchange model and showed that Pareto efficiency and hiding-proofness are incompatible for additive preferences. Moulin (2008) studied splitting and merging in task scheduling problems and provided an impossibility result where any deterministic scheduling procedure is vulnerable to either splitting or merging. A similar concept called false-name manipulations has been considered in several environments (Yokoo, Sakurai, and Matsubara 2004; Aziz et al. 2011; Todo and Conitzer 2013).

## Our Model

There is a set of agents  $N = \{1, \dots, n\}$  and a set of indivisible goods  $K$  in the world. Each agent  $i \in N$  is endowed with a set of goods, or *endowments*,  $w_i \subset K$ . Let  $w = (w_i)_{i \in N}$  be an *endowment distribution* to  $N$ , satisfying  $\bigcup_{i \in N} w_i \subseteq K$  and  $w_i \cap w_j = \emptyset$  for any pair  $i, j \in N$ . The endowment distribution is chosen, by nature, from the set  $\mathcal{W}$  of all possible endowment distributions.

Each agent  $i \in N$  has a linear ordering, or *preference*,  $R_i$ , over the set of all possible bundles  $L \subseteq K$ . Let  $\mathcal{R}$  denote a set of all admissible preferences, or a *preference domain*. Given  $R_i \in \mathcal{R}$  and a pair  $L, L' \subseteq K$ , let  $LR_iL'$  denote that  $L$  is weakly preferred to  $L'$  at  $R_i$ . We assume preferences are *strict*, i.e., for any pair  $L, L' \neq L$ , either  $LP_iL'$  or  $L'P_iL$ , where  $P_i$  indicates the strict component of  $R_i$ . Therefore,  $LR_iL'$  but not  $LP_iL'$  implies  $L = L'$ . We also assume  $\{k\}P_i\emptyset$  for any  $i \in N$ , any  $R_i \in \mathcal{R}$ , and any  $k \in K$ . Let  $R = (R_i)_{i \in N}$  denote a preference profile of  $N$  and  $R_{-i} = (R_j)_{j \neq i}$  denote a preference profile of  $N \setminus \{i\}$ .

In summary, each agent  $i$  has two private information,  $w_i \subset K$  and  $R_i \in \mathcal{R}$ . We refer to  $\theta_i = (w_i, R_i) \in \Theta_{w,i}$  as  $i$ 's *type*, where  $\Theta_{w,i}$  denotes the set of all *reportable types* of  $i$  when an endowment distribution  $w$  is chosen. More precisely, for a given  $w$ , let  $W(w, i)$  denote the set of reportable goods s.t.  $\bigcup_{i \in N} W(w, i) = K$ ,  $W(w, i) \cap W(w, j) = \emptyset$  for any pair  $i, j \in N$ , and  $W(w, i) \supseteq w_i$  for any  $i \in N$ <sup>1</sup>. With this notation,  $\Theta_{w,i} := 2^{W(w,i)} \times \mathcal{R}$ . Note that at this moment there is no restriction on revealed goods  $\hat{w}_i$  of  $i$ , as long as they are in  $W(w, i)$ . Let  $\theta \in \Theta_w = \times_{i \in N} \Theta_{w,i}$  be a type profile of  $N$  and  $\theta_{-i} = (\theta_j)_{j \neq i} \in \Theta_{w,-i}$  be a type profile of agents except  $i$ , where  $\Theta_{w,-i} = \times_{j \neq i} \Theta_{w,j}$ .

We then formally define a set of possible *assignments*. When a set of goods  $L \subseteq K$  is reported by agents, let  $X_L$  be a set of the possible assignments of goods to the agents, s.t. any *assignment*  $x \in X_L$  satisfies  $\bigcup_{i \in N} x_i = L$  and  $x_i \cap x_j = \emptyset$  for any pair  $i, j \in N$ .

An *exchange rule*  $\varphi$  maps a reported type profile  $\hat{\theta} = (\hat{w}_i, \hat{R}_i) \in \Theta_w$  into a possible assignment  $x \in X_{\hat{w}}$ , where  $\hat{w} = \bigcup_{i \in N} \hat{w}_i$ . Here let  $\varphi_i(\hat{\theta})$  be a set of goods assigned to  $i$  by  $\varphi$  when agents report  $\hat{\theta}$ . Note that  $\varphi$  is a function that reallocates all *reported* goods, rather than *exactly owned* goods  $\bigcup_{i \in N} w_i$ , because it could not observe  $w_i$ . Therefore,  $\bigcup_{i \in N} \varphi_i(\hat{\theta}) \supset \bigcup_{i \in N} w_i$  might occur when some  $i$  reports good  $k \in W(w, i) \setminus w_i$ , while this possibility will be eliminated by an assumption in the next section.

We now define some desirable properties considered in this paper. Individual rationality requires that reporting the true preference with revealing true endowments is never worse than not participating.

**Definition 1** (Individual Rationality). *An exchange rule  $\varphi$  is individually rational if  $\forall w \in \mathcal{W}, \forall i \in N, \forall \theta_{-i} \in \Theta_{w,-i}, \forall \theta_i = (w_i, R_i) \in \Theta_{w,i}, \varphi_i(\theta_i, \theta_{-i})R_i w_i$  holds.*

Strategyproofness is an incentive constraint for agents, requiring that reporting a true preference with revealing true

<sup>1</sup>We need this to make each agent’s type space independent from the reports of others.

endowments is weakly better than misreporting her preference or withholding a subset of endowments  $w_i \setminus \hat{w}_i$ .

**Definition 2** (Strategyproofness). *An exchange rule  $\varphi$  is strategyproof if  $\forall w \in \mathcal{W}, \forall i \in N, \forall \theta_{-i} \in \Theta_{w,-i}, \forall \theta_i = (w_i, R_i) \in \Theta_{w,i}, \forall \hat{\theta}_i = (\hat{w}_i, \hat{R}_i) \in \Theta_{w,i}, \varphi_i(\theta_i, \theta_{-i}) R_i \varphi_i(\hat{\theta}_i, \theta_{-i}) \cup (w_i \setminus \hat{w}_i)$  holds.*

Note that  $\varphi_i(\hat{\theta}_i, \theta_{-i}) \cup (w_i \setminus \hat{w}_i)$  indicates the bundle that  $i$  finally owns when she reports  $\hat{\theta}_i$ . In our model, each agent has an option to withhold all her endowments, which is essentially equivalent to not participating. That is, an exchange rule is individually rational if it is strategyproof.

To evaluate the performance of an exchange rule, it is natural to measure the quality, or the efficiency, of assignments that can be produced by it. *Pareto efficiency* is one of the most popular efficiency criteria.

**Definition 3** (Pareto efficiency). *An assignment  $y \in X_L$  Pareto dominates  $x \in X_L$  if  $\forall i \in N, y_i R_i x_i$ , and  $\exists j \in N, y_j P_j x_j$ . An exchange rule  $\varphi$  is Pareto efficient if  $\forall w \in \mathcal{W}, \forall \hat{\theta} \in \Theta_w$ , any  $y \in X_{\hat{w}}$  does not Pareto dominate  $\varphi(\hat{\theta})$ .*

In this paper we focus on four domains of strict preferences. First, we mention that for any  $R_i$ , we can define w.l.o.g. a corresponding utility function  $u_i : 2^K \rightarrow \mathbb{R}$  s.t. for any  $L, L' \subseteq K, LP_i L'$  if and only if  $u_i(L) > u_i(L')$ . Therefore, we define each preference domain,  $\mathcal{R}_m, \mathcal{R}_r, \mathcal{R}_a$ , and  $\mathcal{R}_l$ , as a condition on  $u_i$ . Note that they have inclusion relations of  $\mathcal{R}_m \supseteq \mathcal{R}_r \supseteq \mathcal{R}_a \supseteq \mathcal{R}_l$ .

**Definition 4** (Monotonic Preferences). *A monotonic preference domain  $\mathcal{R}_m$  is the set of all possible preferences  $R_i$  s.t.  $u_i$  satisfies  $\forall L, L' \subseteq K, L \supset L' \Rightarrow u_i(L) > u_i(L')$ .*

**Definition 5** (Responsive Preferences). *A responsive preference domain  $\mathcal{R}_r$  is the set of all possible preferences  $R_i$  s.t.  $u_i$  satisfies  $\forall L \subseteq K, \forall k, k' \in K \cup \{\emptyset\} \setminus L, u_i(k) > u_i(k') \Rightarrow u_i(L \cup \{k\}) > u_i(L \cup \{k'\})$ .*

**Definition 6** (Additive Preferences). *An additive preference domain  $\mathcal{R}_a$  is the set of all possible preferences  $R_i$  s.t.  $u_i$  satisfies  $\exists v_i : K \rightarrow \mathbb{R}$  s.t.,  $\forall L \subseteq K, u_i(L) = \sum_{k \in L} v_i(k)$ .*

**Definition 7** (Lexicographic Preferences). *A lexicographic preference domain  $\mathcal{R}_l$  is the set of all possible preferences  $R_i$  s.t.  $u_i$  satisfies  $\exists v_i : K \rightarrow \{1, 2, \dots, 2^{K-1}\}, \forall L \subseteq K, u_i(L) = \sum_{k \in L} v_i(k)$ , where  $\rightarrow$  indicates a bijection.*

In some real application fields of exchange economies, each agent's utility is sometimes determined almost exclusively by the most preferred good among those she got. For instance, in the housing market, each agent actually uses only one house, while all of the other houses she receives will give it a different amount of utility as real estate. Lexicographic preferences reflect such situations.

When an agent decides an order over singletons, a lexicographic preference is uniquely determined without any additional information. Moreover, if an agent does not have such an order, no full strict preference can be produced due to the lack of information. That is, the lexicographic preference domain is a "smallest" domain of strict preferences. In this point, it is worth discussing exchange problems under lexicographic preference domain, from the perspective of finding a "price of strict preferences."

## The Revelation Principle

The revelation principle guarantees that focusing on strategyproof exchange rules is sufficient, when we are only interested in market with dominant strategy equilibria. However, in our exchange model with private endowments, the revelation principle does not always hold. The following indirect mechanism gives a counter example:

**Example 1.** *Each agent brings her endowments to a market. If someone brings a pre-specified good, say  $\alpha \in K$ , the market runs an (strategyproof) exchange rule to trade goods except for  $\alpha$ , by asking agents for their preferences. Otherwise, do nothing and return all goods to their owners.*

The mechanism has a dominant-strategy equilibrium: each agent first brings all her endowments to the market and, if asked by the mechanism, reports her true preference. However, we cannot truthfully implement the outcome by an exchange rule when agents can report endowments that they do not initially own. E.g., when no agent initially owns such  $\alpha$ , an agent  $i$  s.t.  $\alpha \in W(w, i) \setminus w_i$  would reveal  $\alpha$  as one of her endowments and make the trade processed. This is not strategyproof, and the revelation principle fails.

To fix the failure of the revelation principle, we introduce the following assumption.

**Assumption 1.** *No agent can reveal any good that she does not initially own. Formally,  $\forall w \in \mathcal{W}, \forall i \in N, \forall \theta_i = (w_i, R_i) \in \Theta_{w,i}$ , let  $M(\theta_i) \subseteq \Theta_{w,i}$  be the set of all reportable types  $\hat{\theta}_i = (\hat{w}_i, \hat{R}_i)$ . Then  $\forall \hat{\theta}_i \in M(\theta_i), \hat{w}_i \subseteq w_i$ .*

In such actual markets as eBay and Amazon.com, that assumption is almost satisfied; sellers are often asked to upload photos of their items. Under this policy, it is slightly more difficult for sellers to offer an item that they do not have.

We call such a mapping  $M$  a *misreport restriction system*. When an exchange rule has a misreport restriction system, the strategyproofness condition is drastically weakened:

**Definition 8.** *An exchange rule  $\varphi$  with a misreport restriction system  $M$  is strategyproof if  $\forall w \in \mathcal{W}, \forall i \in N, \forall \theta_{-i} \in \Theta_{w,-i}, \forall \theta_i \in \Theta_{w,i}, \forall \hat{\theta}_i \in M(\theta_i), \varphi_i(\theta_i, \theta_{-i}) R_i \varphi_i(\hat{\theta}_i, \theta_{-i}) \cup (w_i \setminus \hat{w}_i)$  holds.*

Consider an exchange rule  $\varphi$  as direct revelation mechanism, which has a misreport restriction system  $M$  satisfying Assumption 1. Let  $\Gamma = (\mathcal{S}_1, \dots, \mathcal{S}_n, g)$  be an indirect mechanism s.t. each  $i$  has a strategy space  $\mathcal{S}_i(\cdot)$  that varies based on her true type  $\theta_i$ , and goods are assigned to agents by an outcome function  $g : \times_{i \in N} \mathcal{S}_i \rightarrow X$ . In our exchange model, let  $g_i(s)$  be the assignment to  $i$  under the indirect mechanism  $\Gamma$  when agents play a strategy profile  $s \in \times_{i \in N} \mathcal{S}_i$ . We say an indirect mechanism  $\Gamma$  implements an exchange rule  $\varphi$  if there exists a dominant strategy equilibrium  $s^*(\theta) = (s_i^*)_{i \in N}$  s.t.  $\forall w \in \mathcal{W}, \forall \theta \in \Theta_w, \forall i \in N, g_i(s^*(\theta)) = \varphi_i(\theta)$ . The following is the revelation principle for our model with private endowments.

**Theorem 1** (Revelation Principle). *Suppose that an exchange rule  $\varphi$  with a misreport restriction system  $M$  satisfying Assumption 1 is implemented by an indirect mechanism  $\Gamma = (\mathcal{S}_1, \dots, \mathcal{S}_n, g)$  satisfying  $\forall i \in N, \forall \theta_i \in \Theta_{w,i}, \forall \hat{\theta}_i \in M(\theta_i), \mathcal{S}_i(\hat{\theta}_i) \subseteq \mathcal{S}_i(\theta_i)$ . Then  $\varphi$  is strategyproof.*

The proof is omitted since it is straightforward from the original proof of the traditional revelation principle, see e.g. Chapter 23 in (Mas-Colell, Whinston, and Green 1995). In what follows we assume that exchange rules have such misreport restriction system  $M$  satisfying Assumption 1 and thus use the new definition of strategyproofness. Also, we can assume w.l.o.g. that for any  $w \in \mathcal{W}$ ,  $\bigcup_{i \in N} w_i = K$ .

### Impossibility

In this section we show that although Assumption 1 makes the revelation principle hold, no exchange rule satisfies strategyproofness and Pareto efficiency.

**Proposition 1.** *Under the lexicographic preference domain, there exists no exchange rule that satisfies strategyproofness and Pareto efficiency.*

*Proof.* Consider  $N = \{1, 2\}$ ,  $K = \{\alpha, \beta, \gamma\}$ ,  $(w_1, w_2) = (\alpha\beta, \gamma)$ , and  $R_1, R_2$  defined by the following orders over  $K$ :

$$\begin{aligned} R_1 &: \gamma \succ \alpha \succ \beta \\ R_2 &: \alpha \succ \beta \succ \gamma \end{aligned}$$

If both report truthfully, a strategyproof and Pareto efficient rule returns  $(\alpha\gamma, \beta)$ ,  $(\beta\gamma, \alpha)$ , or  $(\gamma, \alpha\beta)$ . If  $(\alpha\gamma, \beta)$  occurs, consider a misreport  $\hat{R}_2 : \alpha \succ \gamma \succ \beta$  by agent 2. Then possible assignments become  $(\beta\gamma, \alpha)$  and  $(\gamma, \alpha\beta)$ , and agent 2 gets higher utility than only receiving  $\beta$  from both assignments. If either  $(\beta\gamma, \alpha)$  or  $(\gamma, \alpha\beta)$  occurs, consider a misreport  $\hat{R}_1 : \alpha \succ \gamma \succ \beta$  by agent 1. The only possible assignment is  $(\alpha\gamma, \beta)$ , which increases agent 1's utility.  $\square$

To show the tightness of the impossibility result, we argue that if we ignore either of these two properties, we can design an exchange rule that satisfies the other property. A dictatorship rule, which gives all goods brought to the market to an agent for any report, is Pareto efficient. However, it obviously violates individual rationality, and thus violates strategyproofness. On the other hand, A no-trade rule, which does not process any trade on the market for any report, is strategyproof but not Pareto efficient.

It might still be worth mentioning that when preferences do not need to be strict, there might be a preference domain under which some rules satisfy both properties, even if ownership is private. For example, when all the goods are identical and the preferences are additive, then the initial endowments are also Pareto efficient. Thus, the no-trade rule satisfies both strategyproofness and Pareto efficiency.

Note that in the proof above, no hiding manipulation is considered, except for deriving individual rationality. Therefore, the result has the following strong implication: even when the ownership of goods is publicly observable (as well as many other works such as Sönmez (1999) and Pápai (2003)), strategyproofness, individual rationality, and Pareto efficiency cannot be satisfied simultaneously under a lexicographic preference domain. In other words, focusing on lexicographic preferences does not help much. Furthermore, as we mentioned before, the lexicographic preference domain is the smallest strict preference domain. This implies that in multiple endowment environments with strict preferences, no exchange rule simultaneously satisfies those properties.

### Strategyproof Exchange Rules

Pápai (2003) introduced the notion of *trading possibility graphs*, which can be regarded as a proposal of trades by an exchange rule to agents. We basically follow the concept, and introduce additional restrictions to guarantee strategyproofness for our model with private ownership.

**Definition 9** (One-for-One Trading Possibility Graph). *Let  $G = (K, E)$  be a directed graph s.t. each node  $k \in K$  corresponds to each good.  $G$  is a one-for-one trading possibility graph if for each node, indegree and outdegree are one.*

Each  $k \in K$  is included in exactly one cycle (if we consider a self-loop a cycle). For given one-for-one trading possibility graph  $G$  and good  $k \in K$ , let  $c(G, k) \subseteq E$  be a cycle containing  $k$ , and let  $k \in c$  mean that cycle  $c$  contains  $k$ . The length of cycle  $c$  is given as the number of nodes in  $c$ , i.e.,  $|c| = \#\{k | k \in c\}$ . Furthermore, let  $C(G)$  be the set of all cycles in  $G$ , i.e.,  $C(G) = \{c(G, k) | k \in K\}$ .

Let us introduce the relationship between revealed goods  $\hat{w}$  and graph  $G$ . For given graph  $G$  and cycle  $c \in C(G)$ , let  $I(G, c)$  be the set of agents in cycle  $c$ , i.e.,  $I(G, c) := \{i \in N | \exists k' \in \hat{w}_i \text{ s.t. } k' \in c\}$ . Also, for given graph  $G$ , cycle  $c \in C(G)$ , and reported goods  $\hat{w}_i$  of agent  $i$ , let  $\hat{w}_{i,c}$  be the set of goods that are revealed by  $i$  and contained in  $c$ , i.e.,  $\hat{w}_{i,c} := \{k \in \hat{w}_i | k \in c\}$ . From the definition of one-for-one trading possibility graph,  $\hat{w}_i = \bigcup_{c \in C(G)} \hat{w}_{i,c}$ .

Let  $p(G, k)$  be the good  $k' \in K$  pointed to by  $k$  in given one-for-one trading possibility graph  $G$ . Formally,  $p(G, k) = k' \text{ s.t. } (k, k') \in E$ . By definition, such a good uniquely exists. For given  $G$  and  $L \subseteq K$ , let  $p(G, L)$  be the set of goods pointed to by  $L$ , i.e.,  $p(G, L) := \bigcup_{k \in L} p(G, k)$ . When its meaning is obvious from the context, we respectively denote  $c(G, k)$ ,  $p(G, k)$ ,  $p(G, L)$ , and  $I(G, c)$  as  $c(k)$ ,  $p(k)$ ,  $p(L)$ , and  $I(c)$  for notation simplicity.

There is one main difference with the original trading possibilities graph proposed by Pápai (2003). Since in our model an exchange rule cannot observe the exact ownership, the one-for-one trading possibility graph cannot always satisfy the requirement of (Pápai 2003) that all goods contained in a cycle are endowed with different agents. The requirement plays an important role in the original paper to guarantee strategyproofness under the responsive preference domain. Therefore, even focusing on responsive preferences, we need to add modifications to Pápai's result (2003).

### Unanimous Agreement Rules

In general resource allocations, a trivial way to guarantee strategyproofness is to recommend a pre-specified solution for them and return it as the outcome when all agents accept it. Utilizing the one-for-one trading possibility graph, we can implement this idea as a class of exchange rules.

**Definition 10** (Unanimous Agreement Rule). *An exchange rule  $\varphi$  is a unanimous agreement rule if there exists a pre-defined one-for-one trading possibility graph  $G$  s.t.  $\forall w \in \mathcal{W}$ ,  $\forall \hat{\theta} = (\hat{w}, \hat{R}) \in \Theta_w$ ,  $\forall i \in N$ ,*

$$\varphi_i(\hat{\theta}) = \begin{cases} p(\hat{w}_i) & \text{if } [\forall k' \in K, p(k') \neq k' \Rightarrow k' \in \hat{w}] \\ & \wedge [\forall i \in N, p(\hat{w}_i) P_i \hat{w}_i] \\ w_i & \text{otherwise.} \end{cases}$$

In words, when there exists a cycle with length of more than one in  $G$ , the rule requires that all goods in the cycle must be revealed by agents; otherwise no trade is done and they receive the *status quo*. Therefore, the range of any unanimous agreement rule is at most two.

For any one-for-one trading possibility graph  $G$ , we can define a unanimous agreement rule. One extreme example in the class is the no-trade rule, which is obtained by setting  $G = (K, E)$  s.t.  $\forall k \in K, (k, k) \in E$ , i.e., no cycle longer than one exists. We can guarantee strategyproofness for any  $G$ , even under the monotonic preference domain.

**Theorem 2.** *Under the monotonic preference domain, any unanimous agreement rule is strategyproof.*

*Proof Sketch.* Intuitively, for each  $i$ , the rule recommends two bundles  $p(w_i)$  and  $w_i$  when it reveals all the endowments. The former is accepted only when  $i$  prefers it under the reported preference  $R_i$ . Therefore, misreports are not beneficial. Even if an agent hides some of her endowments, the rule never recommends any different assignment.  $\square$

### Agreement Cycles Rules

Because of the requirement of a unanimous agreement on a trade proposed by the rule, the range of a unanimous agreement rule contains at most two assignments, causing a much loss of efficiency. To avoid this negative result, we introduce another class of strategyproof exchange rules that utilize the corresponding one-for-one trading possibility graphs in a better way under specific preference domains. Namely, they allow independent decision making for different cycles.

**Definition 11** (Trade via Cycles). *Given one-for-one trading possibility graph  $G$ , cycle  $c \in C(G)$ , set of reported goods  $\hat{w} = \bigcup_{i \in N} \hat{w}_i$ , reported preference profile  $\hat{R}$ , and agent  $i \in N$ , let  $\tau_{i,c}(G, \hat{w}, \hat{R})$  be the set of goods traded with  $\hat{w}_{i,c}$  through cycle  $c$  in the graph  $G$ . More precisely,*

$$\tau_{i,c}(G, \hat{w}, \hat{R}) = \begin{cases} p(\hat{w}_i) & \text{if } [k' \in c \Rightarrow k' \in \hat{w}] \\ & \wedge [\forall j \in I(c), p(\hat{w}_{j,c}) P_j \hat{w}_{j,c}] \\ \hat{w}_i & \text{otherwise.} \end{cases}$$

**Definition 12** (Agreement Cycles Rule). *An exchange rule  $\varphi$  is an agreement cycles rule if there is a pre-defined one-for-one trading possibility graph  $G$  s.t.  $\forall w \in \mathcal{W}, \forall \hat{\theta} = (\hat{w}, \hat{R}) \in \Theta_w, \forall i \in N, \varphi_i(\hat{\theta}) = \bigcup_{c \in C(G)} \tau_{i,c}(G, \hat{w}, \hat{R})$ .*

This is an extension of the unanimous agreement rules. Indeed, when the number of cycles longer than one in  $G$  is less than two, i.e., every vertex pointing to another one is included in one big cycle, then it is also represented as an unanimous agreement rule.

We provide more conditions on one-for-one trading possibility graph  $G$  for the agreement cycles rules to satisfy strategyproofness. We first show that under the additive preference domain, any agreement cycles rule satisfies strategyproofness regardless of the structure of corresponding one-for-one trading possibility graph  $G$ .

**Theorem 3** (Additive Preferences). *Under the additive preference domain, an agreement cycles rule with any one-for-one trading possibility graph is strategyproof.*

*Proof Sketch.* We can observe that under the additive preference domain, an exchange rule is strategyproof if and only if it is robust against each of misreporting preferences and hiding endowments, respectively. From the definition of the agreement cycles rules, a decision on a cycle  $c$  is independently determined from the decisions on other cycles. Since the preferences are additive, it suffices to show that, for each cycle  $c \in C(G)$ , an agent cannot get a preferable bundle by both misreporting preferences and hiding endowments.

Misreporting preferences never raises an agent's utility on a cycle, since the rule only recommends two bundles,  $p(\hat{w}_{i,c})$  and  $\hat{w}_{i,c}$ , to an agent on it, and reporting true preference  $R_i$  always chooses the preferred one. Nor is hiding endowments helpful, since it only reduces the chance that a preferred bundle is chosen on each cycle.  $\square$

It is somewhat obvious that under the additive preference domains, truth-telling is the best choice, since these preferences have no super-additive synergies. However, when agents have such synergies, independent decisions for different cycles will give them the incentive to manipulate. We first show that under the responsive preference domain, truth-telling is not always the best choice when two initial endowments are traded via a cycle for two different goods.

**Theorem 4** (Responsive Preferences). *Under the responsive preference domain, an agreement cycles rule with one-for-one trading possibility graph  $G$  is strategyproof iff*

$$[\nexists c \in C(G) \text{ s.t., } |c| > 3] \vee [\exists c \in C(G) \text{ s.t., } |c| = |K|].$$

*Proof.* We first show the *if* part. When the second condition is satisfied, the rule coincides with a unanimous agreement rule and satisfies strategyproofness. Consider the other case, i.e., the length of each cycle in  $G$  is always at most three.

For given such graph  $G$ , for each cycle  $c \in C(G)$  and for each agent  $i \in I(c)$ , at most one good is traded via cycle  $c$ , even when  $|c| = 3$  and  $|\hat{w}_{i,c}| = 2$ . Because, in that case, one of the two goods is certainly returned to her. Also, at each cycle  $c$ , a truth report  $\theta_i$  gives the agent the better bundle between  $p(c, w_i)$  and  $w_i$  under her true preference.

When we fix decisions on all other cycles  $c' \in C(G)$ , getting the better bundle on  $c$ , which has at most one good different with the other choice on  $c$ , never reduces the agent's utility from responsiveness. By repeatedly applying this, we can see that a truth report is a dominant strategy.

For the *only if* part, consider graph  $G = (K, E)$  s.t.  $K = \{\alpha, \beta, \gamma, \delta, \epsilon\}$  and  $E = \{(\alpha, \alpha), (\beta, \gamma), (\gamma, \delta), (\delta, \epsilon), (\epsilon, \beta)\}$  (Fig. 1 (a)), which violates both conditions above. Note that any one-for-one trading possibility graph violating both of the conditions has the same (or a quite similar) structure as Fig. 1 (a), so focusing on the graph  $G$  is without loss of generality. Let us consider the case where  $N = \{1, 2\}$ ,  $(w_1, w_2) = (\alpha\beta\delta, \gamma\epsilon)$  and  $R_1, R_2 \in \mathcal{R}_r$  satisfy:

$$\begin{aligned} R_1 : & \dots \succ \alpha\beta\delta \succ \dots \succ \alpha\gamma\epsilon \succ \dots \succ \gamma\epsilon \succ \dots \succ \beta\delta \succ \dots \\ R_2 : & \dots \succ \beta\delta \succ \dots \succ \gamma\epsilon \succ \dots \end{aligned}$$

Preference  $R_1$  is not additive, but it is still responsive.

When both agents truthfully report their preferences with revealing all their endowments, a trade via the cycle  $c$  is processed, which gives a bundle  $\alpha\gamma\epsilon$  to agent 1. This violates individual rationality, and thus violates strategyproofness.  $\square$

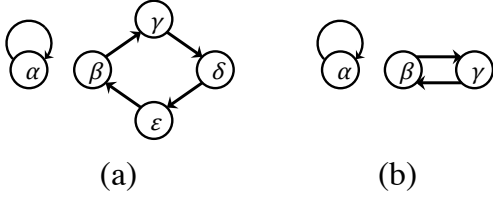


Figure 1: Examples of one-for-one trading possibility graphs violating strategyproofness under  $\mathcal{R}_r$  and  $\mathcal{R}_m$  respectively

In other words, due to the existence of super-additive synergies between a good and a bundle of size more than one, between  $\alpha$  and  $\beta\delta$  in the above example, on a responsive preference, an independent decision on a cycle (without regarding other cycles) might decrease an agent’s utility.

The next example shows that when we consider the monotonic preference domain, any combination of independent decisions could give agents the incentive to manipulate.

**Theorem 5** (Monotonic Preferences). *Under the monotonic preference domain, an agreement cycles rule with one-for-one trading possibility graph  $G$  is strategyproof iff*

$$[\forall c \in C(G) \text{ s.t., } |c| > 1] \vee [\exists c \in C(G) \text{ s.t., } |c| = |K|].$$

*Proof.* The *if* part is proven from the fact that when  $G$  satisfies either of the conditions, the rule coincides with a unanimous agreement rule with identical graph  $G$ .

To show the *only if* part, consider graph  $G = (K, E)$  s.t.  $K = \{\alpha, \beta, \gamma\}$  and  $E = \{(\alpha, \alpha), (\beta, \gamma), (\gamma, \beta)\}$  (see Fig. 1 (b)), which violates both above conditions. Note that any one-for-one trading possibility graph violating both conditions has the same structure as Fig. 1 (b), so focusing on the graph  $G$  is without loss of generality. Consider  $N = \{1, 2\}$ ,  $(w_1, w_2) = (\alpha\beta, \gamma)$ , and  $R_1, R_2 \in \mathcal{R}_m$  satisfy:

$$\begin{aligned} R_1 : & \dots \succ \alpha\beta \succ \alpha\gamma \succ \dots \succ \gamma \succ \beta \succ \dots \\ R_2 : & \dots \succ \beta \succ \gamma \succ \dots \end{aligned}$$

Preference  $R_1$  is not responsive, but it is still monotone.

When both truthfully report their preferences with revealing all their endowments, a trade via the cycle proceeds, which gives a bundle  $\alpha\gamma$  to agent 1. This violates individual rationality and thus violates strategyproofness. Almost the same discussion appears in Pápai (2007).  $\square$

## Robustness to Splitting

One of the real application fields in which the ownership of goods is private is an exchange economy on the Internet. In such environments, however, since the possibilities of different kinds of manipulations have been pointed out, guaranteeing strategyproofness is sometimes not sufficient. For instance, an agent may *split* her endowments into two bundles and participate in a market under two identities, one of which is fake, e.g., under different email accounts. In this section we show that our proposed exchange rules are also *split-proof*, meaning that no agent has an incentive to split her endowments into multiple bundles or misreports her preference whenever they are strategyproof.

Note that even under the monotonic preference domain, the unanimous agreement rules are not the only strategyproof ones. For example, consider the following rule: first separate agents into parties, and propose for each party a trade (or a deal) *based on the reports of agents outside of the party*. This is obviously strategyproof, but it cannot be represented as any unanimous agreement rule. Pápai (2003, 2007) designed strategyproof rules that utilize the knowledge of the ownership of goods to propose trades to agents.

However, when ownership is private and agents might create fakes, there is almost no way to propose trades without being affected by agents’ strategies. Indeed, the above example is not robust against splitting (Moulin 2008); there is a chance that an agent’s fake is in a different party from her true identity. Such an agent might benefit by misreporting under the fake to change a proposal to her original party.

The following is the main result of this section. We omit the formal definition of split-proofness and a detailed proof due to space limitations, but we believe readers can easily understand the underlying intuition.

**Theorem 6.** *Any unanimous agreement rule is split-proof under the monotonic preference domain. Also, any agreement cycles rule is split-proof under any preference domain under which it is strategyproof.*

*Proof Sketch.* Any unanimous agreement rule is obviously split-proof, since for each agent, a possible assignment is either her initial endowments or the one proposed by the one-for-one trading possibility graph, regardless of her splitting.

This is also true for each cycle in agreement cycles rules. Furthermore, since each agent originally has the right to make independent decisions on different cycles via misreporting preferences, splitting is not more powerful than only misreporting preferences. Thus, an agreement cycles rule is split-proof whenever it is strategyproof.  $\square$

## Conclusions

In this paper we studied the effect of private ownership in exchange economies with multiple goods. We presented two fundamental results on strategyproof exchange rules; the revelation principle holds under a natural assumption, and even under the lexicographic preference domain, Pareto efficiency and strategyproofness are incompatible. We then proposed agreement cycles rules, which perform slightly better than trivial strategyproof rules under more restricted domains. All of these rules are also split-proof in a stronger sense that split manipulations are assumed to be costless.

Our exchange model remains quite simple due to the lack of indifferences/incomparability/unacceptability in the preferences. Extending it to such richer preferences is a natural research direction. It seems also interesting to design randomized rules (Ashlagi et al. 2010) with higher social welfare. Finally, even though we focused on strategyproofness, we are also interested in studying Pareto efficient rules that are *computationally hard* to manipulate (Pini et al. 2011).

## Acknowledgement

This work was partially supported by JSPS KAKENHI Grant Number 24220003.

## References

- Ashlagi, I.; Fischer, F. A.; Kash, I. A.; and Procaccia, A. D. 2010. Mix and match. In *ACM Conference on Electronic Commerce*, 305–314.
- Atlamaz, M., and Klaus, B. 2007. Manipulation via endowments in exchange markets with indivisible goods. *Social Choice and Welfare* 28(1):1–18.
- Aziz, H.; Bachrach, Y.; Elkind, E.; and Paterson, M. 2011. False-name manipulations in weighted voting games. *J. Artif. Int. Res.* 40(1):57–93.
- Chen, Y., and Sönmez, T. 2002. Improving efficiency of on-campus housing: An experimental study. *American Economic Review* 92(5):1669–1686.
- Ma, J. 1994. Strategy-proofness and the strict core in a market with indivisibilities. *International Journal of Game Theory* 23(1):75–83.
- Mas-Colell, A.; Whinston, M.; and Green, J. 1995. *Microeconomic Theory*. Oxford student edition. Oxford University Press.
- Moulin, H. 2008. Proportional scheduling, split-proofness, and merge-proofness. *Games and Economic Behavior* 63(2):567–587.
- Pápai, S. 2000. Strategyproof assignment by hierarchical exchange. *Econometrica* 68(6):1403–1433.
- Pápai, S. 2003. Strategyproof exchange of indivisible goods. *Journal of Mathematical Economics* 39(8):931 – 959.
- Pápai, S. 2007. Exchange in a general market with indivisible goods. *Journal of Economic Theory* 132(1):208 – 235.
- Pini, M. S.; Rossi, F.; Venable, K. B.; and Walsh, T. 2011. Manipulation complexity and gender neutrality in stable marriage procedures. *Autonomous Agents and Multi-Agent Systems* 22(1):183–199.
- Postlewaite, A. 1979. Manipulation via endowments. *Review of Economic Studies* 46(2):255–262.
- Roth, A. E., and Postlewaite, A. 1977. Weak versus strong domination in a market with indivisible goods. *Journal of Mathematical Economics* 4(2):131–137.
- Roth, A. E.; Sönmez, T.; and Ünver, M. U. 2004. Kidney exchange. *The Quarterly Journal of Economics* 119(2):457–488.
- Shapley, L., and Scarf, H. 1974. On cores and indivisibility. *Journal of Mathematical Economics* 1(1):23–37.
- Sönmez, T. 1999. Strategy-proofness and essentially single-valued cores. *Econometrica* 67(3):677–689.
- Todo, T., and Conitzer, V. 2013. False-name-proof matching. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems, AAMAS '13*, 311–318. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.
- Yokoo, M.; Sakurai, Y.; and Matsubara, S. 2004. The effect of false-name bids in combinatorial auctions: New fraud in internet auctions. *Games and Economic Behavior* 46(1):174–188.
- Yu, L. 2011. Mechanism design with partial verification and revelation principle. *Autonomous Agents and Multi-Agent Systems* 22(1):217–223.