Complexity consideration on the existence of strategy-proof social choice functions

Koji TAKAMIYA*

Economics, Niigata University

July 13, 2010

Abstract

Social choice theorists have long recognized that in models of private goods economies, strategy-proofness is sometimes incompatible with individual rationality plus Pareto efficiency, and that it is usually more or less "difficult" to prove this incompatibility. In this paper we examine this "difficulty" from the viewpoint of computational complexity. We set up a simple model of private goods exchange where agents bring in and trade indivisible objects under consumption constraints. We consider the computational problem of deciding whether for a given specification of the economy, there exists a social choice function which is strategy-proof, individually rational and Pareto efficient. We prove that (i) this is an $\mathcal{NP}$-hard problem, and point out, however, that (ii) the problem becomes computationally trivial if we drop one of these three properties of the social choice function.

JEL Classification— C72, C78, D71, D78.

Keywords— computational complexity; allocation of indivisible objects; $\mathcal{NP}$-hardness; strategy-proofness; individual rationality; Pareto efficiency.

1 Introduction

In the traditional literature of social choice, it has been a central issue to investigate the *existence* of social choice procedures which satisfy various desirable properties from the viewpoints of incentive, efficiency, equity and so on. Among many themes in this realm, the existence of *strategy-proof* social choice functions has attracted significant attention for many years. It has been long recognized that strategy-proofness often conflicts with other desirable properties. And not only that there are conflicts but also it is often

*Faculty of Economics, Niigata University. 8050 2-no-cho Ikarashi Nishi-ku Niigata 950-2181 JAPAN. tel: +81-25-262-6498. E-mail: takamiya@econ.niigata-u.ac.jp.

difficult to establish that there *is* indeed a conflict, i.e. to prove that strategy-proofness is incompatible with some other desirable properties. For example, the celebrated Gibbard-Satterthwaite theorem (Gibbard 1973, Satterthwaite 1975) depicts the conflict between strategy-proof and *non-dictatorship*, a very weak requirement of equity. And this conflict was difficult to establish: The Gibbard-Satterthwaite theorem had been conjectured many years before it was proved. It took a long time for the theorem to be proved.

More recent studies on strategy-proof functions in models of *private goods economies* have revealed the conflict between strategy-proofness and *individual rationality plus Pareto efficiency*, which is also often difficult to establish. For example, in 1972, Hurwicz proved that for any classical pure exchange economy with *two persons* and *two goods*, if the preference domain includes a sufficiently wide set of classical preferences, then there does not exist a social choice function which is strategy-proof, individually rational and Pareto efficient (Hurwicz, 1972). He conjectured that the same result holds true for those economies with three or more agents and goods. However, this problem had remained unsolved for about thirty years until Serizawa's work appeared (Serizawa, 2002), which proved that Hurwicz's result is generalized to the case of any finite numbers of agents and goods.

Another example is from the theory of matching models such as of the *marriage problem* (Gale and Shapley, 1962) and the *housing market* (Shapley and Scarf, 1972). It had been known from the early 1980's that (i) for the marriage problem with the full strict preference domain[1], no *core stable rule* (i.e. a social choice function which chooses a core stable matching for each preference profile) is strategy-proof, and that (ii) in contrast, for the housing market also with the full strict preference domain, there exists the unique core stable rule and this rule is strategy-proof (Roth, 1982a; Roth 1982b). (Note that any core stable rule is both individually rational and Pareto efficient.) Clearly these two results exhibit a sharp contrast. However, for a long time, it had not been fully understood where this sharp contrast came from. It was 1999 when Sönmez provided an answer to this question: He set up a general model of indivisible objects allocation, which covers both the marriage problem and the housing market, and proved the following: Provided that the preference domain is the full strict preference domain, if a social choice function is strategy-proof, individually rational and Pareto efficient, then it must be that for each preference profile, the core (i.e. the set of core stable allocations) is a singleton unless it is empty, and that this function chooses the core stable allocation whenever available. In the marriage problem the core is neither a singleton nor empty for some preference profile. Thus Sönmez's result implies the nonexistence of strategy-proof functions which are individaully raiotional and Pareto efficient in the marriage problem. Later, Takamiya (2003) showed a conditional converse of Sönmez's result: Provided that

[1]The *full strict preference domain* is the domain where each agent can have any strict ranking over the agent's own assignments (i.e. no consumption externalities). This domain is usually assumed in social choice analysis of matching problems.

the preference domain is the full strict preference domain, if the core is a singleton for each preference profile, then the unique core stable rule is strategy-proof. Evidently the strategy-proofness of the core stable rule in the housing market immediately follows from this result. These two results of Sönmez and Takamiya have provided some understanding in the existence problem of strategy-proof functions by relating it to the singletonness of the core. However, to this date, it has not been fully investigated under what conditions the singletonness of the core is obtained in the general setting formulated by Sönmez. This seems to be a hard combinatorial problem.[2]

To date, social choice theorists know from experience (partially described as above) that in models of private goods economies, it is usually more or less difficult to decide whether there exists a social choice function which satisfies these three properties altogether. This is in contrast to that it is also known that it is usually easy to obtain strategy-proof functions which are individually rational *or* Pareto efficient *separately.* For example, in most models of private goods economies, it is trivial to have a strategy-proof function which is Pareto efficient only: A *dictatorial* function, in which some fixed agent always receives all the goods in the economy, is both strategy-proof and Pareto efficient.

The purpose of the present research is to examine the idea that *in private goods economies, it is difficult to determine whether there exists a social choice function which is strategy-proof, individually rational and Pareto efficient.* Our approach is metaphorical in the sense that we do not directly analyze those problems which social choice theorists have attacked or do not go into the ingenuity of their proofs. Rather, for our analytical purpose, we set up a simple and artificial problem and analyze its difficulty of a specific kind: To embody the concept of "difficulty" we employ the concept of *time complexity* in the theory of *computational complexity.*

Concretely, our analysis is as follows: We give a simple model of private goods economies where agents bring in and trade indivisible objects. There each agent is faced with a consumption constraint. This model is a special case of the general model of indivisible objects allocation formulated in Sönmez's above-mentioned paper. We consider the computational problem of deciding whether for a given specification of the economy (i.e. a given instance of the problem), there exists a social choice function which is strategy-proof, individually rational and Pareto efficient. First, for our main theorem, we prove that this is an $\mathcal{NP}$-hard problem. Here $\mathcal{NP}$-hardness captures the idea of "difficulty" in deciding the existence of such functions. Second, we point out, however, that this problem becomes computationally *trivial* if we drop one of these three properties of the social choice function. That is, for any two properties out of these three properties, for

[2]This problem has been partially solved: In the setting of the *coalition formation problem*, a special case of Sönmez's general model, Pápai (2004) has provided a necessary and sufficient condition for the core to be a singleton for each preference profile in the full strict preference domain. However, to our knowledge, the computational complexity of checking this condition has not been investigated.

any instance of the problem there exists at least one social choice function which satisfies these two properties. Thus the answer to the decision problem is always "yes".

It is important to note that the nature of our research is different from that of most lines of research in computational social choice. Usually in computational social choice theory, computational ideas are used to formulate and analyze various realistic constraints put on the prosecution of social choice procedures or the behavior of agents who act in the social choice process, which arise from the limited availability of material and mental resources. However, here computational ideas are employed to express the idea of the difficulty which (traditional) social choice researchers face with when they look for desirable strategy-proof functions. In this sense, our paper is still research in computational social choice but more precisely is "research about the traditional social choice research from the viewpoint of computation".

2 Preliminaries

2.1 Economic Model

Let us define the economic model that we examine. We consider modeling reallocation of multiple indivisible objects. An **allocation problem** is a list $\mathcal{E} = (N, \Omega, \{\Theta^i\}_{i\in N}, u, (w, q), x_0)$. Here N is the set of **agents**, and Ω is the set of **(indivisible) objects**. N and Ω are both assumed to be nonempty finite sets.

An **allocation** is a set-valued function $x : N \rightarrow\rightarrow \Omega$ which is "partitional," i.e. (i) $i \neq j \Longrightarrow x(i) \cap x(j) = \emptyset$ and (ii) $\bigcup_N x(i) = \Omega$. Let us denote the set of allocations by $\mathcal{X}$.

$u := \{u^i\}_{i\in N}$ are **utility functions**, and Θ^i is the **type space** of agent i. Each element of Θ^i specifies the preference of agent i in the following way: A **value function** $v : (\Theta^1 \cup \ldots \cup \Theta^n) \times \Omega \to \mathbb{Z}$ is defined. We assume utility functions are all *additive* with respect to the values given by the value function. That is, the utility function of agent i, $u^i : \mathcal{X} \times \Theta^i \to \mathbb{Z}$ is defined so as to satisfy the following: for all $x \in \mathcal{X}$ and $\theta^i \in \Theta^i$,

$$u^i(x, \theta^i) = \sum_{\omega\in x(i)} v(\theta^i, \omega). \tag{1}$$

Note that values $v(\theta^i, \cdot)$ could be negative and so are utility levels.

(w, q) is a **feasibility constraint**, which consists of **weights** w and **capacities** q. Here w is a function $w : N \times \Omega \to \mathbb{Z}_+$, and q is a function $q : N \to \mathbb{Z}_+$. Here $\mathbb{Z}_+ := \{0, 1, 2, \cdots\}$. For agent $i \in N$, object $\omega \in \Omega$ has a weight $w(i, \omega)$, and i can consume a bundle of objects unless the sum of the weights of these objects exceeds i's capacity $q(i)$. Thus it is defined that an allocation $x \in \mathcal{X}$ is **feasible** to agent $i \in N$ if

$$\sum_{\omega\in x(i)} w(i, \omega) \leq q(i). \tag{2}$$

An allocation $x \in \mathcal{X}$ is called **feasible** if it is feasible to all the agents. Let us denote the set of feasible allocations by $\mathcal{X}^f$. In the following we refer to feasible allocations simply as allocations.

Finally, x_0 denotes the **initial endowments**. We assume $x_0 \in \mathcal{X}^f$.

Let $x \in \mathcal{X}^f$ and $\theta \in \Theta$. Then x is **individually rational** at θ if for any $i \in N$,

$$u^i(x, \theta^i) \geq u^i(x_0, \theta^i). \tag{3}$$

And x is **Pareto efficient** at θ if for any $y \in \mathcal{X}^f$

$$[\forall i \in N,\ u^i(y, \theta^i) \geq u^i(x, \theta^i)] \Longrightarrow [\forall i \in N,\ u^i(y, \theta^i) = u^i(x, \theta^i)]. \tag{4}$$

2.2 Relevance of the model

Our model is a special case of the general allocation model studied by Sönmez (1999) (which we have mentioned in Sec.1). To Sönmez's model, we have added specific structures on admissible preferences and feasible allocations, namely, *additivity* of utilities, *weights* of objects, *capacities* of agents. These structures admit concise representations of feasible allocations and preferences. Without such structures, inputs of the problem can be overly redundant, which apparently reduces the complexity of the problem.

We admit that as a model of private goods economies, our model is unusual and perhaps artificial in assuming weights and capacities. However, it is still relevant as a modeling of economic problems. For example, our model includes the housing market (Shapley and Scarf, 1974), an important economic model, as a special case. Further, in some cases, we may interpret weights as *personalized prices* of objects and capacities as *budgets* that agents face.

2.3 Properties of social choice functions

Let an allocation problem be given. Let us denote $\Theta := \Theta^1 \times \Theta^2 \times \cdots \times \Theta^n$. Any element θ of Θ is called a **type profile**. A **social choice function** is a function $f : \Theta \to \mathcal{X}^f$. We consider the following three properties of social choice functions.

- **Strategy-proofness.** Let $i \in N$ and $\theta \in \Theta$. Then we say that i **manipulates** f at θ if for some $\tilde{\theta}^i \in \Theta^i$,

$$u^i(f(\theta^{-i}, \tilde{\theta}^i), \theta^i) > u^i(f(\theta^{-i}, \theta^i), \theta^i), \tag{5}$$

  f is called **strategy-proof** if for any $i \in N$, i cannot manipulate f at any $\theta \in \Theta$.

- **Individual rationality.** Let us call f **individually rational** if for any $\theta \in \Theta$, x is individually rational at θ.

- **Pareto efficiency.** Let us call f **Pareto efficient** if for any $\theta \in \Theta$, x is Pareto efficient at θ.

3 Results

3.1 Main theorem

We consider the decision problem in the following. Let a positive integer $\bar{n}$ be given.

NAME: $\mathrm{SP} + \mathrm{IR} + \mathrm{PE}(\bar{n})$
INSTANCE: An allocation problem $\mathcal{E} = (N, \Omega, \{\Theta^i\}_{i \in N}, u, (w, q), x_0)$ with $|N| = \bar{n}$.
QUESTION: Does there exist a social choice function for $\mathcal{E}$ which is strategy-proof, individually rational and Pareto efficient?

Our main result is that $\mathrm{SP} + \mathrm{IR} + \mathrm{PE}(\bar{n})$ is *$\mathcal{NP}$-hard*. Note that in our formulation of the computational problem above, the number of agents is fixed, i.e. $|N| = \bar{n}$. Without this restriction on the number of agents, it is not much surprising even if the problem would be computationally hard because the space of type profiles grows exponentially as the number of agents gets larger.

Theorem 1 *The problem* $\mathrm{SP} + \mathrm{IR} + \mathrm{PE}(\bar{n})$ *is $\mathcal{NP}$-hard if $\bar{n} \geq 4$.*

The construction made in our proof of Theorem 1 requires four agents at least ($\bar{n} \geq 4$). We do not know whether $\mathrm{SP} + \mathrm{IR} + \mathrm{PE}(\bar{n})$ is $\mathcal{NP}$-hard if $\bar{n} = 2$ or 3.

3.2 Interpretation of the main theorem

(1) To understand the subtlety of Theorem 1, we should be aware of the following fact.

Theorem 2 *Let an allocation problem $\mathcal{E} = (N, \Omega, \{\Theta^i\}_{i \in N}, u, (w, q), x_0)$ with an arbitrary size of $|N|$ be given. And let us pick up any two of the three properties, strategy-proofness, individual rationality and Pareto efficiency. Then there exists a social choice function which satisfies these two properties.*

Theorem 2 says that if we drop one of the three properties of social choice functions which are listed in the problem $\mathrm{SP} + \mathrm{IR} + \mathrm{PE}(\bar{n})$, then the computational problem becomes *trivial*: The answer is "yes" for *any instance*. This fact tells us that what makes the computational problem hard is neither each single requirement of strategy-proofness, individual rationality or Pareto efficiency, or even each *pair* of these three properties, but rather is the *combination* of these *three* properties altogether.

(2) It is important to notice that what is at issue here is the computational problem deciding the *existence* of social choice functions which satisfy some properties. One should carefully distinguish this problem from the computational problem of deciding whether a given social choice function satisfies those properties. In fact, the latter problem can be computationally very hard without combining these three properties. For example, if we

are given an allocation problem and a type profile is fixed, then the problem of deciding whether a given allocation is not Pareto efficient is $\mathcal{NP}$-complete. To state more precisely, the following theorem holds. Let us define the following problem: Let a natural number $\bar{n}$ be given.

NAME: NOTPARETO($\bar{n}$)
INSTANCE: An allocation problem $\mathcal{E} = (N, \Omega, \{\Theta^i\}_{i\in N}, u, (w, q), x_0)$ with $|N| = \bar{n}$ and $\Theta = \{\theta\}$; and an allocation $x \in \mathcal{X}^f$.
QUESTION: Is x not Pareto efficient?

Theorem 3 *The problem* NOTPARETO($\bar{n}$) *is $\mathcal{NP}$-complete if $\bar{n} \geq 2$.*

From the above theorem, it directly follows that the problem of deciding whether a given social choice function is not Pareto efficient is also $\mathcal{NP}$-complete. On the contrary, it is computationally trivial to decide whether a Pareto efficient social choice function *exists* because such a function always exists.

Our Theorem 3 follows from Theorem 1 in the paper of de Keijzer, Bouveret, Klos and Zhang (2009), which studies computational problems arising from an allocation model of indivisible objects with additive utilities.[3] However, in Sec.3.3, we will give our own proof of Theorem 3, which utilizes a construction used in our proof of Theorem 1.

3.3 Proofs

For the preparation of proving Theorem 1, let us consider the following allocation problem $\mathcal{E}_1$.

- $N = \{1, 2, 3\}$.
- $\Omega = \{c_1, c_2, c_3\}$.
- $x_0^1 = \{c_1\}$; $\quad x_0^2 = \{c_2, c_3\}$; $\quad x_0^3 = \emptyset$.
- $\Theta^1 = \{\theta_1^1, \theta_2^1, \theta_3^1, \theta_4^1, \theta_5^1, \theta_6^1\}$; $\quad \Theta^2 = \{\theta_1^2, \theta_2^2, \theta_3^2, \theta_4^2, \theta_5^2, \theta_6^2\}$; $\quad \Theta^3 = \{\theta^3\}$.
- The following table depicts the values of $v(\theta_j^1, c_k)$, $v(\theta_j^2, c_k)$ and $v(\theta^3, c_k)$.

Table 1.

	θ_1^1	θ_2^1	θ_3^1	θ_4^1	θ_5^1	θ_6^1	θ_1^2	θ_2^2	θ_3^2	θ_4^2	θ_5^2	θ_6^2	θ^3
$v(\cdot, c_1)$	1	1	2	2	3	3	3	3	2	2	1	1	0
$v(\cdot, c_2)$	2	3	1	3	1	2	2	1	3	1	3	2	0
$v(\cdot, c_3)$	3	2	3	1	2	1	1	2	1	3	2	3	0

[3]We are thankful to an anonymous referee for notifying us of the work of de Keijzer *et al.*

- The following table depicts the values of $w(i, c_k)$ and $q(i)$.

Table 2.

	1	2	3
$w(\cdot, c_1)$	1	1	0
$w(\cdot, c_2)$	1	1	0
$w(\cdot, c_3)$	1	1	0
$q(\cdot)$	1	2	0

Lemma 1 *For the allocation problem $\mathcal{E}_1$, there does not exist any social choice function that is strategy-proof, individually rational and Pareto efficient.*

Proof. Suppose that f is strategy-proof, individually rational and Pareto efficient. Let us denote allocations $x \in \mathcal{X}^f$ by 3-tuples, i.e. $x = (x(1), x(2), x(3))$. Let

$$x_1 = (\{c_1\}, \{c_2, c_3\}, \emptyset), \quad x_2 = (\{c_2\}, \{c_3, c_1\}, \emptyset), \quad x_3 = (\{c_3\}, \{c_1, c_2\}, \emptyset). \tag{6}$$

Since f is Pareto efficient, for all $\theta \in \Theta$, $f(\theta) \in \{x_1, \ x_2, \ x_3\}$. Given the above, clearly, for each of agents 1 and 2, the agent's possible preferences can be regarded as the set of strict rankings over $\{x_1, \ x_2, \ x_3\}$. Further, we can ignore the existence of agent 3. Therefore, f is regarded as a social choice function with three alternatives and two agents whose admissible preferences are exactly the set of strict rankings of the three alternatives. Then since f is strategy-proof and Pareto efficient, by the Gibbard-Satterthwaite theorem (Gibbard 1973, Satterthwaite 1975) f is dictatorial, i.e. there exists some $i \in \{1, 2\}$ such that for any $\theta \in \Theta$, $f(\theta)$ equals the allocation that maximizes agent i's utility at θ^i. Clearly f violates individual rationality. Thus we reach the desired conclusion. □

Proof of Theorem 1. Clearly it suffices to prove only for the case where $\bar{n} = 4$ because one can increase the number of agents by adding dummy agents who does not have any initial assignments and is not able to receive any objects for the capacity constraint. The proof is done by reduction from the following problem PARTITION ([SP 12] in Gary and Johnson (1979)).

NAME: PARTITION
INSTANCE: A finite set $A = \{a_1, a_2, \ldots, a_p\}$ and a function $s : A \to \mathbb{N}$.
QUESTION: Does there exist a partition $\{A_1, A_2\}$ of A such that $\sum_{a \in A_1} s(a) = \sum_{a \in A_2} s(a)$.

Let an instance of PARTITION (A, s) be given. Then we give a polynomial-time transformation of this instance into an instance of $\mathrm{SP} + \mathrm{IR} + \mathrm{PE}(4)$ in the following, and we will show that the answer to this instance of PARTITION is “yes” if and only if the the answer to this instance of SP + IR + PE(4) is “yes”.

Let us consider the following instance $\mathcal{E}_2$ of SP + IR + PE(4). We denote $\sum_{a \in A} s(a)$ by $s(A)$ in the sequel.

- $N = \{1, 2, 3, 4\}$.
- $\Omega = A \cup \{b, c_1, c_2, c_3\}$.
- $x_0^1 = \{c_1\}$; $x_0^2 = \{c_2, c_3\}$; $x_0^3 = A$; $x_0^4 = \{b\}$.
- $\Theta^1 = \{\theta_1^1, \theta_2^1, \theta_3^1, \theta_4^1, \theta_5^1, \theta_6^1\}$; $\Theta^2 = \{\theta_1^2, \theta_2^2, \theta_3^2, \theta_4^2, \theta_5^2, \theta_6^2\}$; $\Theta^3 = \{\theta^3\}$; $\Theta^4 = \{\theta^4\}$.
- The following table depicts the values of $v(\cdot, \cdot)$.

Table 3.

	θ_1^1	θ_2^1	θ_3^1	θ_4^1	θ_5^1	θ_6^1	θ_1^2	θ_2^2	θ_3^2	θ_4^2	θ_5^2	θ_6^2	θ^3	θ^4
$v(\cdot, c_1)$	1	1	2	2	3	3	3	3	2	2	1	1	0	0
$v(\cdot, c_2)$	2	3	1	3	1	2	2	1	3	1	3	2	0	0
$v(\cdot, c_3)$	3	2	3	1	2	1	1	2	1	3	2	3	0	0
$v(\cdot, a_1)$	0	0	0	0	0	0	0	0	0	0	0	0	0	$2s(a_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$v(\cdot, a_i)$	0	0	0	0	0	0	0	0	0	0	0	0	0	$2s(a_i)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$v(\cdot, a_p)$	0	0	0	0	0	0	0	0	0	0	0	0	0	$2s(a_p)$
$v(\cdot, b)$	4	4	4	4	4	4	0	0	0	0	0	0	0	$s(A) - 1$

- The following table depicts the values of $w(\cdot, \cdot)$ and $q(\cdot)$.

Table 4.

	1	2	3	4
$w(\cdot, c_1)$	1	1	0	$s(A) + 1$
$w(\cdot, c_2)$	1	1	0	$s(A) + 1$
$w(\cdot, c_3)$	1	1	0	$s(A) + 1$
$w(\cdot, a_1)$	2	3	0	$2s(a_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$w(\cdot, a_i)$	2	3	0	$2s(a_i)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$w(\cdot, a_p)$	2	3	0	$2s(a_p)$
$w(\cdot, b)$	1	3	1	$s(A)$
$q(\cdot)$	1	2	0	$s(A)$

Note the two key facts of this construction: (a) Agent 3 brings the objects $A = \{a_1, a_2, \ldots, a_p\}$ into the economy as the initial allocation, and these objects are valuable only to agent 4. (b) If agent 4 leaves this economy with his initial assignment and agent 3's initial assignment is deleted, then the resulting economy is identical with $\mathcal{E}_1$ constructed for Lemma 1.

(i) First, we show that if the answer to the PARTITION instance (A, s) is "yes", then that to the $\mathrm{SP} + \mathrm{IR} + \mathrm{PE}(4)$ instance constructed above is "yes". Now suppose that the answer to the PARTITION instance is "yes". In the following we show that for any type profile, there is only one allocation which is both individually rational and Pareto efficient, and that the social choice function which chooses this allocation for any type profile (thus individually rational and Pareto efficient) is strategy-proof.

Let us fix a type profile. Given the date above, it is clear that agent 4 is to be better off by releasing the object b and instead collecting some objects A' out of A if and only if A' satisfies $\sum_{a \in A'} 2s(a) = s(A)$. And that the answer to the PARTITION instance (A, s) is "yes" means that such A' exists. This reallocation (agent 4 releases b and obtains A') is Pareto improvement because the objects A are valuable only to agent 4 and chucking out the object b never hurts the other agents' utility. And this fills up agent 4's capacity. Further, Pareto efficiency forces the object b to go to agent 1, and this fills up agent 1's capacity. By Pareto efficiency, agent 2 receives the two his most preferred objects out of $\{c_1, c_2, c_3\}$ depending on agent 2's type, and this fills up agent 2's capacity. Finally agent 3 receives the remaining object from $\{c_1, c_2, c_3\}$ and A. Obviously this allocation is individually rational. This is the unique allocation which is individually rational and Pareto efficient.

Let us consider the social choice function which chooses the unique individually rational and Pareto efficient allocation for each type profile. Agents 1 and 4 receives the same assignment for any type profile, and only the assignments of agents 2 and 3 vary. Agent 2 obtains his most preferred assignment. And agent 3's utility level is constant whatever this agent receives. Therefore, there is no situation where some agent can manipulate the outcome, that is, this function is strategy-proof.

(ii) Second, we show that if the answer to the PARTITION instance (A, s) is "no", then that to the $\mathrm{SP} + \mathrm{IR} + \mathrm{PE}(4)$ instance is "no". Now suppose that the answer to the PARTITION instance is "no". In this case, it is not possible for agent 4 to improve his utility level by receiving some objects from A in return for giving up the object b. Thus individual rationality forces agent 4 to keep the object b that fills up agent 4's capacity. Now for the feasibility constraint, any object in A cannot go to either agents 1 or 2 so agent 3 has to keep all the objects of A. Therefore, by the fact (b) indicated above, now the situation is identical with the economy $\mathcal{E}_1$. Then by applying Lemma 1, we conclude that there does not exist any strategy-proof social choice function which is individually rational and Pareto efficient. □

Proof of Theorem 2. (i) There exists a social choice function which is *individually rational* and *Pareto efficient*. Because it is clear that for every type profile, there exists at least one allocation which is both Pareto efficient and individually rational.

(ii) There exists a social choice function which is *strategy-proof* and *individually rational*. An example is the *constant* function, which always chooses the initial endowments x_0.

(iii) There exists a social choice function which is *strategy-proof* and *Pareto efficient*. A social choice function based on *serial dictatorship* (Satterthwaite and Sonnenschein, 1981) satisfies both properties. In the following we define this class of functions and prove that any function in this class satisfies these two properties: Let π be a bijection from $\{1, 2, \ldots, |N|\}$ to N. For each $\theta \in \Theta$, the sets $C^\pi(\theta, i)$ $(i = 0, 1, 2, \ldots, |N|)$ is defined inductively as follows:

$$C^\pi(\theta, 0) = \mathcal{X}^f, \tag{7}$$

$$\text{for } i = 1, 2, \ldots, |N|, \quad C^\pi(\theta, i) = \arg \max_{x \in C^\pi(\theta,\ i-1)} u^{\pi(i)}(x, \theta^{\pi(i)}). \tag{8}$$

Note that if $i < j$, then $C^\pi(\theta, j) \subset C^\pi(\theta, i)$. A social choice function f is a **serial dictatorship based on** π if for all $\theta \in \Theta$, $f(\theta) \in C^\pi(\theta, |N|)$.

First, we show that for any bijection $\pi : \{1, 2, \ldots, |N|\} \to N$, any serial dictatorship f on π is Pareto efficient: Let $x \in f(\theta)$ and $y \in \mathcal{X}^f$. Suppose $\forall i \in N,\ u^i(y, \theta^i) \geq u^i(x, \theta^i)$. Then, first of all, we have $y \in C^\pi(\theta, 1)$ because otherwise it would be $u^{\pi(1)}(y, \theta^{\pi(1)}) < u^{\pi(1)}(x, \theta^{\pi(1)})$, a contradiction. Next, we note that for any $i \in \{2, 3, \ldots, |N|\}$ if $y \in C^\pi(\theta, i-1)$, then $y \in C^\pi(\theta, i)$ because otherwise it would be also $u^{\pi(i)}(y, \theta^{\pi(i)}) < u^{\pi(i)}(x, \theta^{\pi(i)})$. Consequently, by induction, we have $y \in C^\pi(\theta, |N|)$, which implies $\forall i \in N,\ u^i(y, \theta^i) = u^i(x, \theta^i)$. Thus we conclude that x is Pareto efficient.

Second, it is easy to see that these f are also strategy-proof. Because of the way serial dictatorship is defined, any agent $\pi(i)$ receives one of his best assignments among $C^\pi(\theta, i-1)$. However, $C^\pi(\theta, i-1)$ is fully determined by $(\theta^{\pi(1)}, \ldots, \theta^{\pi(i-1)})$ so agent $\pi(i)$'s reporting of his type does not affect this set. Thus $\pi(i)$ cannot be better off by misreporting his type. □

Proof of Theorem 3. First, we show that NOTPARETO(2) (so NOTPARETO($\bar{n}$) with $\bar{n} \geq 2$) is $\mathcal{NP}$*-hard* by reduction from PARTITION. Let an instance (A, s) of PARTITION be given. Let us give a polynomial-time transformation of this instance into an instance of NOTPARETO(2) as follows. The following construction is based on the same idea as the gadget consisting of agents 3 and 4 in the proof of Theorem 1 above.

- $N = \{1, 2\}$.
- $\Omega = A \cup \{b\}$.

- $x_0^1 = A; \quad x_0^2 = \{b\}$
- The following table depicts the values of v.

Table 5.

	θ^1	θ^2
$v(\cdot, a_1)$	0	$2s(a_1)$
$\vdots$	$\vdots$	$\vdots$
$v(\cdot, a_i)$	0	$2s(a_i)$
$\vdots$	$\vdots$	$\vdots$
$v(\cdot, a_p)$	0	$2s(a_p)$
$v(\cdot, b)$	0	$s(A) - 1$

- The following table depicts the values of w, q.

Table 6.

	1	2
$w(\cdot, a_1)$	0	$2s(a_1)$
$\vdots$	$\vdots$	$\vdots$
$w(\cdot, a_i)$	0	$2s(a_i)$
$\vdots$	$\vdots$	$\vdots$
$w(\cdot, a_p)$	0	$2s(a_p)$
$w(\cdot, b)$	0	$s(A)$
$q(\cdot)$	0	$s(A)$

- The allocation x equals x_0.

Now the allocation x is not Pareto efficient if and only if there exists a subset A' of A such that $\sum_{a \in A'} 2s(a) = s(A)$. (Because if such A' exists, agent 2 can be better off without hurting agent 1's utility by releasing the object b and instead collecting A'.) And that the answer to the PARTITION instance (A, s) is "yes" if and only if such A' exists. This establishes the $\mathcal{NP}$-hardness of NOTPARETO.

Second, it is easy to see NOTPARETO$(\bar{n}) \in \mathcal{NP}$. If the answer to a NOTPARETO$(\bar{n})$ instance is "yes" i.e. the considered allocation x is *not* Pareto efficient, then there exists some other allocation y which Pareto dominates x. Now y is a *certificate* and it can be checked in polynomial time whether y Pareto dominates x. $\square$

Acknowledgments

This research is supported by a Grant-in-Aid for Scientific Research of the Ministry of Education, Culture, Sports, Science and Technology of Japan, No.21730155. I am grateful to Ryuichiro Ishikawa, Nobuyuki Hanaki, Toyotaka Sakai, Tomoichi Shinotsuka, Takuma Wakayama and Naoki Watanabe for various comments, and in particular to Takashi Ui for his invaluable critical comments and to Akira Tanaka for many conversations on various topics in computer science and economics. I also thank participants of the seminars at Tsukuba University and Yokohama National University. And I am very thankful to two anonymous referees of COMSOC 2010 for their comments which have improved this paper. All errors are my own responsibility.

References

[1] de Keijzer B, Bouveret S, Klos T and Zhang Y (2009): On the complexity of efficiency and envy-freeness in fair division of indivisible goods with additive preferences, *Proceedings of the First International Conference on Algorithmic Decision Theory*: 98–110.

[2] Hurwicz L (1972): On informationally decentralized systems, In *Decision and Organization.* G. B. McGuire and R. Radner (Eds.) North Holland. Amsterdam.

[3] Gibbard A (1973): Manipulation of voting schemes: A general result, *Econometrica* 41: 587–601.

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

[5] Garey MR and Johnson DS (1979): *Computers and Intractability: A Guide to the Theory of NP-completeness*. WH Freeman. New York.

[6] Pápai S (2004): Unique stability in simple coalition formation games, *Games and Economic Behavior* 48: 337–54.

[7] Roth AE (1982a) The economics of matching: Stability and incentives, *Mathematics of Operations Research* 7: 617–28.

[8] Roth AE (1982b): Incentive compatibility in a market with indivisible goods, *Economics Letters* 9: 127–32.

[9] Satterthwaite MA (1975): Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions, *Journal of Economic Theory* 10: 187–217.

[10] Satterthwaite MA and Sonnenschein H (1981): Strategy-proof allocation mechanisms at differentiable points, *Review of Economic Studies* 48: 587–97.

[11] Serizawa S (2002): Inefficiency of strategy-proof rules for pure exchange economies, *Journal of Economic Theory* 106: 219–241.

[12] Shapley L and Scarf H (1974): On cores and indivisibility, *Journal of Mathematical Economics* 1: 23–37.

[13] Sönmez T (1999): Strategy-proofness and essentially single-valued cores, *Econometrica* 67: 677–89.

[14] Takamiya K (2003): On strategy-proofness and essentially single-valued cores: A converse result, *Social Choice and Welfare* 20: 77–83.