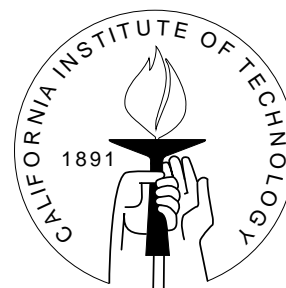# SECURE IMPLEMENTATION: STRATEGY-PROOF MECHANISMS RECONSIDERED

Tatsuyoshi Saijo
Osaka University
and
California Institute of Technology

Tomas Sjöström
The Pennsylvania State University

Takehiko Yamato
Tokyo Institute of Technology

Secure Implementation:
Strategy-Proof Mechanisms Reconsidered[†]

Tatsuyoshi Saijo*

Tomas Sjöström**

and

Takehiko Yamato***

June, 1998/Revised September, 2003

* Institute of Social and Economic Research, Osaka University, Ibaraki, Osaka 567-0047, Japan, E-mail address: saijo@iser.osaka-u.ac.jp; Research Institute of Economy, Trade and Industry, 1-3-1 Kasumigaseki, Chiyoda, Tokyo 100-8901, Japan; and Research Institute of Market Structure, 1-38-9 Nihonbashi Kakigara, Chuo, Tokyo 103-0014, Japan.

** Department of Economics, The Pennsylvania State University, Kern Graduate Building, University Park, PA 16802-3306, USA. E-mail: jts10@psu.edu

*** Department of Value and Decision Science, Graduate School of Decision Science and Technology, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan. E-mail address: yamato@valdes.titech.ac.jp

ABSTRACT

Strategy-proofness, requiring that truth-telling is a dominant strategy, is a standard concept in social choice theory. However, the concept of strategy-proofness has serious drawbacks. First, announcing one's true preference may not be a unique dominant strategy, and using the wrong dominant strategy may lead to the wrong outcome. Second, almost all strategy-proof mechanisms have a continuum of Nash equilibria, and most of which produce the wrong outcome. Third, experimental evidence shows that most of the strategy-proof mechanisms do not work well. We argue that a possible solution to this dilemma is to require double implementation in Nash equilibrium and in dominant strategies, which we call secure implementation. We characterize environments where secure implementation is possible, and compare it with dominant strategy implementation. An interesting example of secure implementation is a Groves mechanism when preferences are single-peaked.

Journal of Economic Literature Classification Number: C92, D71, D78, and H41.

Correspondent:

Tatsuyoshi Saijo
Institute of Social and Economic Research
Osaka University
Ibaraki, Osaka 567-0047
Japan
Phone 81-6 (country & area codes)
6879-8582 (office)/6878-2766 (fax)
E-mail: saijo@iser.osaka-u.ac.jp

1. Introduction

Strategy-proofness, requiring that truth-telling is a dominant strategy, is a standard concept in social choice theory. Although it seems natural to assume that an agent will tell the truth if it is a dominant strategy to do so, there are some complications. First, announcing one's true preference may not be a *unique* dominant strategy, and using the wrong dominant strategy may lead to the wrong outcome. Second, many strategy-proof mechanisms have a continuum of Nash equilibria, and most of which produce the wrong outcome. Third, experimental evidence shows that most of the strategy-proof mechanisms do not work well, that is, very few subjects reveal their true valuations. For example, see Attiyeh, Franciosi, and Isaac (2000) and Kawagoe and Mori (2001) for pivotal mechanism experiments, and Kagel, Harstad, and Levin (1987) and Kagel and Levin (1993) for second price auction experiments with independent private values.

The first problem can be solved by requiring "full" implementation in dominant strategies. That is, all dominant strategy equilibria should yield a socially optimal outcome. In order to cope with the second problem, we provide a new concept called *secure* implementation. A social choice function is securely implementable if there exists a game form or a mechanism that simultaneously implements the social choice function in dominant strategy equilibria and in Nash equilibria. Thus, all Nash equilibria should yield a socially optimal outcome. We characterize securely implementable social choice functions: a social choice function is securely implementable if and only if it satisfies strategy-proofness and a new property called the *rectangular property*. The question of whether secure mechanisms work well in experiments is investigated in a companion paper (Cason, Saijo, Sjöström and Yamato (2003)).

Most strategy-proof mechanisms do not satisfy the rectangular property. For example, the pivotal mechanism for public projects and the serial cost sharing

1

mechanism for an excludable public good have a continuum of Nash equilibria. However, all Groves mechanisms with single-peaked preferences do satisfy the rectangular property. Since secure implementation is a more demanding concept than anything previously proposed in the literature, it is not surprising that the rectangular property is rarely satisfied. We consider secure implementation to be a benchmark: if secure mechanisms do not work well in experiments, then there is very little hope that anything will work. But if a secure mechanism works well in experiments while implementation using less demanding equilibrium concepts fail, then we may be able to pinpoint the reason for the failure by comparing with the benchmark of secure implementation.

The first person to study the relationship between dominant strategy implementation and Nash implementation was Repullo (1985). His main result is that if some social choice function f is dominant strategy implemented by some indirect mechanism, but f is not dominant strategy implemented by its associated direct mechanism, then the indirect mechanism does not Nash implement f. He concluded that the concept of dominant strategy implementation should be discarded in favor of Nash or Bayesian Nash implementation, and "the only role of dominant strategies would be that of ensuring the existence of direct mechanisms that implement the social choice rules under consideration in Nash or Bayesian Nash strategies" (Repullo (1985), p. 229). We agree that the existence of "bad" Nash equilibria is problematic. However, the experimental literature suggests that mechanisms designed for Nash (or Bayesian Nash) implementation may not necessarily work well. In the absence of a dominant strategy, a player's best response depends on the other players' choices, which may be hard to predict. This strategic uncertainty may lead to the failure to coordinate on a Nash (or Bayesian Nash) equilibrium. Thus, neither of the standard concepts – dominant strategy implementation and Nash implementation – provides a robust foundation for practical implementation. However, if a mechanism simultaneously

implements a social choice function in dominant strategies *and* in Nash equilibria, then we get the advantages of dominant strategies (strategic uncertainty is not important), but we avoid the possibility that the players may play "bad" Nash equilibria. In this paper, we investigate the consequences of this strong requirement of secure implementation.

The remainder of the paper is organized as follows. We give notation and definitions in Section 2. In Section 3 we show that many well-known strategy-proof mechanisms have continuum Nash equilibria. We characterize secure implementability in Section 4. Section 5 is for applications of our characterization result to public good economies with quasi-linear preferences. In Section 6 we discuss the relationship between dominant strategy implementation and secure implementation. Concluding remarks are in Section 7.

2. Notation and Definitions

Let A be an arbitrary set of alternatives, and let $I = \{1, 2, ..., n\}$ be the set of agents, with generic element i. We assume that $n \geq 2$. Each agent i is characterized by a preference relation defined over A. We assume that agent i's preference relations admit a numerical representation $u_i : A \to \Re$. For each i, let U_i be the class of utility functions admissible for agent i. Let $u = (u_1, ..., u_n) \in U \equiv \times_{i \in I} U_i$.

A *social choice function* (SCF) is a function $f : U \to A$ that associates with every $u \in U$ a unique alternative $f(u)$ in A.

A *mechanism* (or *game form*) is a function $g : S \to A$ that assigns to every $s \in S$ a unique element of A, where $S = \times_{i \in I} S_i$, S_i is the *strategy space of agent i*. The list $s \in S$ will be written as (s_i, s_{-i}), where $s_{-i} = (s_1, ..., s_{i-1}, s_{i+1}, ..., s_n) \in S_{-i} \equiv \times_{j \neq i} S_j$. Given $s \in S$ and $s_i' \in S_i$, (s_i', s_{-i}) is the list $(s_1, ..., s_{i-1}, s_i', s_{i+1}, ..., s_n)$ obtained by replacing the i-th component of s by s_i'. Let $g(S_i, s_{-i})$ be the *attainable set of agent i at* s_{-i}, i.e., the set of outcomes that agent i can induce when the other agents select s_{-i}.

For $i \in I$, $u_i \in U_i$, and $a \in A$, let $L(a, u_i) \equiv \{b \in A \mid u_i(a) \ge u_i(b)\}$ be the *weak lower contour set for agent i with u_i at a.* Given a mechanism $g: S \to A$, the strategy profile $s \in S$ is a *Nash equilibrium of g at $u \in U$* if for all $i \in I$, $g(S_i, s_{-i}) \subseteq L(g(s), u_i)$. Let $N^g(u)$ be the set of Nash equilibria of g at u. Also, let $N_A^g(u)$ be the set of Nash equilibrium outcomes of g at u, i.e., $N_A^g(u) \equiv \{a \in A \mid$ there exists $s \in S$ such that $s \in N^g(u)$ and $g(s) = a\}$. The mechanism g *implements the SCF f in Nash equilibria* if for all $u \in U$, $f(u) = N_A^g(u)$. f is *Nash implementable* if there exists a mechanism which implements f in Nash equilibria.

Let a mechanism $g: S \to A$ be given. The strategy $s_i \in S_i$ is a *dominant strategy for agent $i \in I$ of g at $u_i \in U_i$* if for all $\hat{s}_{-i} \in S_{-i}$, $g(S_i, \hat{s}_{-i}) \subseteq L(g(s_i, \hat{s}_{-i}), u_i)$. Let $DS_i^g(u_i)$ be the set of dominant strategies for i of g at u_i. The strategy profile $s \in S$ is a *dominant strategy equilibrium of g at $u \in U$* if for all $i \in I$, $s_i \in DS_i^g(u_i)$. Let $DS^g(u)$ be the set of dominant strategy equilibria of g at u. Also, let $DS_A^g(u)$ be the set of dominant strategy equilibrium outcomes of g at u, i.e., $DS_A^g(u) \equiv \{a \in A \mid$ there exists $s \in S$ such that $s \in DS^g(u)$ and $g(s) = a\}$. The mechanism g *implements the SCF f in dominant strategy equilibria* if for all $u \in U$, $f(u) = DS_A^g(u)$. f is *dominant strategy implementable* if there exists a mechanism which implements f in dominant strategy equilibria.

The SCF f is *strategy-proof* if for all $i \in I$, for all $u_i, \tilde{u}_i \in U_i$, for all $\tilde{u}_{-i} \in U_{-i}$, $u_i(f(u_i, \tilde{u}_{-i})) \ge u_i(f(\tilde{u}_i, \tilde{u}_{-i}))$. The following result is well-known:

Proposition 1 *(The Revelation Principle for Dominant Strategy Implementation. Gibbard (1973)). If the SCF f is dominant strategy implementable, then f is strategy-proof.*

The converse of Proposition 1 is not true: some strategy-proof SCF's cannot be dominant strategy implemented (e.g., Dasgupta, Hammond, and Maskin (1979)).

3. The Trouble with Strategy-Proof Mechanisms

In this section, we consider several strategy-proof mechanisms which have been extensively studied in the literature: the pivotal mechanism for a non-excludable public good, the serial cost sharing mechanism for an excludable public good, the second price auction for an indivisible good, the Condorcet winner voting scheme (a median voter scheme) with single-peaked preferences, and the uniform allocation rule (a fixed-price trading rule) with single-peaked preferences. We will show that each of these strategy-proof mechanisms may have a continuum of Nash equilibria. Moreover, there may exist "bad" Nash equilibria outcomes that are Pareto-inferior to the dominant strategy equilibrium outcome. This might explain why many strategy-proof mechanisms do not work well in experiments and they are not used in real economic situations.

(1) The pivotal mechanism (Clarke (1971)).

Consider a two-agent economy with a binary non-excludable public good and quasi-linear preferences. Two agents 1 and 2 are facing a decision whether they should produce the public good or not. Agent i's true net value of the public good is v_i if it is produced, and her true net value is 0 otherwise ($i = 1,2$). In the pivotal mechanism, each agent i reports his net value $\tilde{v}_i$ and the outcome is determined as follows:

Rule 1: if $\tilde{v}_1 + \tilde{v}_2 \geq 0,$ then the public good is produced, and if not, then it is not produced; and

Rule 2: each agent i must pay the pivotal tax t_i

$$t_i = \begin{cases} 0 & \text{if (i) } \tilde{v}_j(\tilde{v}_1 + \tilde{v}_2) > 0 \text{ or (ii) } \tilde{v}_j > 0 \text{ and } \tilde{v}_1 + \tilde{v}_2 = 0 \\ |\tilde{v}_j| & \text{otherwise} \end{cases}$$

where $j \neq i$.

First, let $(v_1, v_2) = (5, -4)$ be the true net value vector. Figure 1-(a) shows that the set of Nash equilibria is approximately a half of the two dimensional area. Notice that

5

the public good should be produced because the sum of the net values of the public good is positive. The upper-right part of the set of Nash equilibria is good since constructing the public good is recommended. However, the lower-left part of the set of Nash equilibria is *bad* since producing the public good is not recommended.

Second, let $(v_1, v_2) = (5,5)$ be the true net value vector. In this case, both agents want to construct the public good. However, Figure 1-(b) shows the area of bad Nash equilibria is still large.

Figure 1 is around here.

We will generalize the above negative result with a binary public good and two agents to the case with any arbitrary finite numbers of public projects and agents.

(2) The serial cost sharing mechanism (Moulin (1994))

Consider a two-agent economy with a binary excludable public good and quasi-linear preferences. The cost of producing the public good is fixed and it is c. Let $(\tilde{v}_1, \tilde{v}_2)$ be a reported gross value vector, and let c_i be the cost that agent i must pay for. Then the serial cost sharing mechanism is defined by the following four rules assuming that $\tilde{v}_1 \geq \tilde{v}_2 \geq 0$:

Rule 1: if $\tilde{v}_1 + \tilde{v}_2 < c$, then the public good is not built;

Rule 2: if $\tilde{v}_1 + \tilde{v}_2 \geq c$, $\tilde{v}_1 < c$, and $\tilde{v}_2 < c/2$, then it is not built;

Rule 3: if $\tilde{v}_1 + \tilde{v}_2 \geq c$, $\tilde{v}_1$, $\tilde{v}_2 \geq c/2$, then both agents enjoy the public good, and

$c_1 = c_2 = c/2$; and

Rule 4: if $\tilde{v}_1 + \tilde{v}_2 \geq c$, $\tilde{v}_1 \geq c$, $\tilde{v}_2 < c/2$, then only agent 1 enjoys the public good, and

$c_1 = c$.

Let $(v_1, v_2) = (5,4)$ be the true gross value vector and $c = 6$ be the cost. Figure 2 shows that the set of Nash equilibria is quite large and the set of bad Nash equilibria cannot be ignored.

<div align="center">

Figure 2 is around here.

</div>

(3) The second price auction (Vickrey (1961)).

Consider a two-agent model with an indivisible good. Two agents 1 and 2 are facing a decision who receives the indivisible good. Agent i's true value of the good is $v_i \geq 0$ if she receives it, and her true value is 0 otherwise ($i = 1,2$). Let $(\tilde{v}_1, \tilde{v}_2)$ be a reported value vector. The second price auction consists of two rules:

Rule 1: if $\tilde{v}_i > \tilde{v}_j$, then agent i receives the good and pays $\tilde{v}_j$ ($i, j = 1,2; i \neq j$); and

Rule 2: if $\tilde{v}_1 = \tilde{v}_2$, then agent 1 receives the good and pays $\tilde{v}_2$.

Let $(v_1, v_2) = (7, 5)$ be the true value vector. Figure 3 shows that the set of Nash equilibria is quite large. Notice that agent 1 should receive the good because her value is greater than agent 2's. The lower-right part of the set of Nash equilibria is good since agent 1 receives the good. However, the upper-left part of the set of Nash equilibria is bad since agent 2 receives the good.

<div align="center">

Figure 3 is around here.

</div>

(4) The Condorcet winner voting scheme (a median voter scheme) with single-peaked preferences.

Consider a voting model with three agents and three alternatives, $N = \{1,2,3\}$ and $A = \{a_1, a_2, a_3\}$. Each agent i's preferences are single-peaked and the most preferred alternative (peak) according to her true preferences is a_i:

<div align="center">7</div>

$u_1(a_1) > u_1(a_2) > u_1(a_3)$, $u_2(a_2) > u_2(a_1)$, $u_2(a_2) > u_2(a_3)$, and $u_3(a_3) > u_3(a_2) > u_3(a_1)$.

In the Condorcet winner voting scheme, each agent i reports her most preferred alternative (peak) $\widetilde{a}_i \in \{a_1, a_2, a_3\}$, and the outcome is determined as follows:

Rule 1: if at least two agents report the same alternative, then it is chosen; and

Rule 2: otherwise, a_2 is chosen.

As Figure 4 illustrates, there exist five Nash equilibria other than the dominant strategy equilibrium of the true peak reporting. Now suppose that utilities are transferable and symmetric: $u_1(a_1) = u_2(a_2) = u_3(a_3)$, $u_1(a_2) = u_2(a_1) = u_2(a_3) = u_3(a_2)$, and $u_1(a_3) = u_3(a_1)$. Then only the outcome a_2 is efficient because the sum of utilities is maximized. The three Nash equilibria are good in the sense that the efficient outcome a_2 is chosen. However, there are the two "bad" Nash equilibria in which the other inefficient outcomes are chosen.

Figure 4 is around here.

A similar problem arises when the set of alternatives is an interval (Moulin (1980)).[1] In fact, it is easy to check that the set of the Nash equilibrium outcomes is equal to the whole set of alternatives,[2] and all Nash equilibrium outcomes other than one dominant strategy equilibrium outcome (the median peak) are bad with transferable and symmetric utilities.

---

[1] Strategy-proof voting schemes have been studied with more general structures of the set of alternatives, including an interval as a special case. For example, see Border and Jordan (1983), Barberà, Massó, and Serizawa (1998), and Schummer and Vohra (2001).

[2] For any point a in the interval, there exists a Nash equilibrium whose outcome is a: each agent reports a.

(5) The uniform rule with single-peaked preferences (a fixed price trading rule) (Sprumont (1991)).[3]

There are two agents, 1 and 2, who must share one unit of some divisible good. Each agent i 's preference relations are single-peaked. Let the peak of the preference relation $\tilde{u}_i$ be $p(\tilde{u}_i) \in [0,1]$. In the uniform rule, each agent i reports $\tilde{u}_i$ and the allocation $(f_1(\tilde{u}), f_2(\tilde{u}))$ is determined as follows:

Rule 1: $f_i(\tilde{u}) = \min\{p(\tilde{u}_i), \lambda(\tilde{u})\}$ if $p(\tilde{u}_1) + p(\tilde{u}_2) \geq 1$, and

Rule 2: $f_i(\tilde{u}) = \max\{p(\tilde{u}_i), \mu(\tilde{u})\}$ if $p(\tilde{u}_1) + p(\tilde{u}_2) \leq 1$ $(i = 1,2)$,

where $\lambda(\tilde{u})$ solves the equation $\min\{p(\tilde{u}_1), \lambda(\tilde{u})\} + \min\{p(\tilde{u}_2), \lambda(\tilde{u})\} = 1$ and $\mu(\tilde{u})$ solves the equation $\max\{p(\tilde{u}_1), \mu(\tilde{u})\} + \max\{p(\tilde{u}_2), \mu(\tilde{u})\} = 1$.

Let $p_i = p(u_i)$ be the peak of the true preference relation u_i. Suppose that $p_2 < 0.5 < p_1$. Figure 5 (a)-(c) illustrate the set of Nash equilibria of this rule in terms of peaks of reported preferences, $\tilde{p}_i = p(\tilde{u}_i)$. Suppose $p_1 + p_2 = 1$ (see Figure 5-(a)). Then the Nash equilibria in the lower-right part are good, since the efficient allocation (p_1, p_2) is assigned. However, there are many bad Nash equilibria in the upper-left part in which inefficient allocations are assigned.[4] Similar things hold for the case with $p_1 + p_2 < 1$ (Figure 5-(b)) as well as for the case with $p_1 + p_2 > 1$ (Figure 5-(c)). On the other hand, if (a) $p_1 > 0.5$ and $p_2 > 0.5$, or (b) $p_1 < 0.5$ and $p_2 < 0.5$, then there exist Nash equilibria other than the truth-telling dominant strategy equilibrium, but all Nash equilibria are good (see Figures 5-(a) and (b)).

Figure 5 and 6 are around here.
----

[3] The uniform rule can be regarded as fixed price trading due to Barberà and Jackson (1995) with two agents, two private goods, and a fixed price ratio.

[4] For example, the equilibrium allocations are $(\tilde{p}_1, \tilde{p}_2)$ if $0.5 \leq \tilde{p}_1 < p_1$ and $p_2 < \tilde{p}_2 \leq 0.5$; and $(0.5, 0.5)$ if either (i) $\tilde{p}_1 = 0.5$ and $0.5 \leq \tilde{p}_2 \leq 1$ or (ii) $0 \leq \tilde{p}_1 \leq 0.5$ and $\tilde{p}_2 = 0.5$. These are Pareto inferior to (p_1, p_2).

4. Secure Implementation: A Characterization and a Revelation Principle

In the previous section, we saw that many strategy-proof mechanisms may have "bad" Nash equilibria. In order to overcome this difficulty, we introduce the following new concept of implementation.

Definition 1. The mechanism g *securely implements the SCF f* if for all $u \in U$, $f(u) = DS_A^g(u) = N_A^g(u)$.[5] The SCF f is *securely implementable* if there exists a mechanism which securely implements f.

Secure implementation requires that for every possible preference profile, the f-optimal outcome equals the set of dominant strategy equilibrium outcomes as well as with the set of Nash equilibrium outcomes.

Next we characterize the class of securely implementable SCF's. We use two conditions. The first condition is strategy-proofness. As Proposition 1 indicates, strategy-proofness is necessary for dominant strategy implementation, and so it is also necessary for secure implementation. However, an additional condition is also necessary for secure implementation. To see why intuitively, suppose that the direct revelation mechanism $g = f$ securely implements the SCF f. See Figure 7 in which $n = 2$ and (u_1, u_2) is the true preference profile.

Suppose that

(4.1) $u_1(f(u_1, \tilde{u}_2)) = u_1(f(\tilde{u}_1, \tilde{u}_2))$,

that is, agent 1 is indifferent between reporting the true preference u_1 and reporting another preference $\tilde{u}_1$ when agent 2's report is $\tilde{u}_2$. Since reporting u_1 is a dominant strategy by strategy-proofness, it follows from (1) that

---

[5] Secure implementation is identical with *double* implementation in dominant strategy equilibria and Nash equilibria. It was Maskin (1979) who first introduced the concept of double implementation. See also Yamato (1993). Note that secure implementation can be regarded as multiple (more than double) implementation in dominant strategy equilibria, Nash equilibria, and all refinements of Nash equilibria whose sets are larger than the set of dominant strategy equilibria.

$u_1(f(\widetilde{u}_1,\widetilde{u}_2)) = u_1(f(u_1,\widetilde{u}_2)) \geq u_1(f(u_1',\widetilde{u}_2))$ for all $u_1' \in U_1$, that is, reporting $\widetilde{u}_1$ is one of agent 1's best responses when agent 2 reports $\widetilde{u}_2$.

Next suppose that

(4.2)　$u_2(f(\widetilde{u}_1,u_2)) = u_2(f(\widetilde{u}_1,\widetilde{u}_2))$.

By using an argument similar to the above, it is easy to see that

$u_2(f(\widetilde{u}_1,\widetilde{u}_2)) \geq u_2(f(\widetilde{u}_1,u_2'))$ for all $u_2' \in U_1$, that is, reporting $\widetilde{u}_2$ is one of agent 2's best responses when agent 1 reports $\widetilde{u}_1$. Therefore, $f(\widetilde{u}_1,\widetilde{u}_2)$ is the Nash equilibrium outcome. Moreover, $f(u_1,u_2)$ is the dominant strategy outcome, and by secure implementability, the dominant strategy outcome coincides with the Nash equilibrium outcome. Accordingly we conclude that $f(u_1,u_2) = f(\widetilde{u}_1,\widetilde{u}_2)$ if (4.1) and (4.2) holds.

Figure 7 is around here.

A formal definition of this condition, called the *rectangular property*, is given as follows:

Definition 3. The SCF f satisfies the *rectangular property* if for all $u,\widetilde{u} \in U$, if $u_i(f(\widetilde{u}_i,\widetilde{u}_{-i})) = u_i(f(u_i,\widetilde{u}_{-i}))$ for all $i \in I$, then $f(\widetilde{u}) = f(u)$.

A formal proof of the claim that the rectangular property is necessary for sure implementation is given as follows:

Lemma 1. *If the SCF f is securely implementable, then f satisfies the rectangular property.*

Proof of Lemma 1: Let $g: S \to A$ be a mechanism which securely implements f. Take any $u,\widetilde{u} \in U$. Suppose that

(4.3)　$u_i(f(\widetilde{u}_i,\widetilde{u}_{-i})) = u_i(f(u_i,\widetilde{u}_{-i}))$ for all $i \in I$.

11

Choose a dominant strategy profile at $\widetilde{u}$, $s(\widetilde{u}) = (s_1(\widetilde{u}_1),...,s_n(\widetilde{u}_n)) \in DS^g(\widetilde{u})$. By dominant implementability,

(4.4) $g(s_1(\widetilde{u}_1),...,s_n(\widetilde{u}_n)) = f(\widetilde{u})$.

Let $i \in I$ be given. Choose a dominant strategy for i at u_i, $s_i(u_i) \in DS_i^g(u_i)$. Then $(s_i(u_i), s_{-i}(\widetilde{u}_{-i})) \in DS^g(u_i, \widetilde{u}_{-i})$, where $s_{-i}(\widetilde{u}_{-i}) = (s_j(\widetilde{u}_j))_{j \neq i}$. By dominant implementability,

(4.5) $g(s_i(u_i), s_{-i}(\widetilde{u}_{-i})) = f(u_i, \widetilde{u}_{-i})$.

By (4.3), (4.4), and (4.5),

(4.6) $u_i(g(s_i(u_i), s_{-i}(\widetilde{u}_{-i}))) = u_i(g(s_1(\widetilde{u}_1),...,s_n(\widetilde{u}_n)))$.

Further, since $s_i(u_i) \in DS_i^g(u_i)$,

(4.7) $g(S_i, s_{-i}(\widetilde{u}_{-i})) \subseteq L(g(s_i(u_i), s_{-i}(\widetilde{u}_{-i})), u_i)$.

By (4.6) and (4.7), $g(S_i, s_{-i}(\widetilde{u}_{-i})) \subseteq L(g(s_i(\widetilde{u}_i), s_{-i}(\widetilde{u}_{-i})), u_i)$. Since this holds for any $i \in I$, $(s_1(\widetilde{u}_1),...,s_n(\widetilde{u}_n)) \in N^g(u)$. By Nash implementability and (4.4), $f(u) = g(s_1(\widetilde{u}_1),...,s_n(\widetilde{u}_n)) = f(\widetilde{u})$. Q.E.D.

The mechanism g is called the *direct revelation mechanism associated with the SCF f* if $S_i = U_i$ for all $i \in I$ and $g(u) = f(u)$ for all $u \in U$. Strategy-proofness and the rectangular property are not only necessary, but also sufficient for secure implementability by the direct revelation mechanism.

Lemma 2. *If the SCF f satisfies strategy-proofness and the rectangular property, then the direct revelation mechanism associated with f securely implements f.*

Proof of Lemma 2: First we prove that for all $u \in U$, $f(u) \in DS_A^g(u)$. Take any $u \in U$. Let $s_i = u_i$ for each $i \in I$. By strategy-proofness, $s \in DS^g(u)$ and $g(s) = f(u)$.

Next we prove that for all $u \in U$, $N_A^g(u) = f(u)$. Let $u \in U$ be given. Take any $s = \widetilde{u} \in N^g(u)$. We show that $g(s) = f(u)$, i.e., $f(\widetilde{u}) = f(u)$. Since $\widetilde{u} \in N^g(u)$,

(4.8) $u_i(f(\tilde{u}_i, \tilde{u}_{-i})) \geq u_i(f(\hat{u}_i, \tilde{u}_{-i}))$ for all $i \in I$ and all $\hat{u}_i \in U_i$.

Further, since $u_i \in DS_i^g(u_i)$ by strategy-proofness,

(4.9) $u_i(f(u_i, \tilde{u}_{-i})) \geq u_i(f(\hat{u}_i, \tilde{u}_{-i}))$ for all $i \in I$ and all $\hat{u}_i \in U_i$.

By (4.8) and (4.9), $u_i(f(\tilde{u}_i, \tilde{u}_{-i})) = u_i(f(u_i, \tilde{u}_{-i}))$ for all $i \in I$. By the rectangular property, $f(\tilde{u}) = f(u)$. Q.E.D.

By Proposition 1, Lemmata 1 and 2, we have the following characterization of securely implementable SCF's.

Theorem 1. *An SCF is securely implementable if and only if it satisfies strategy-proofness and the rectangular property.*

It is easy to check that the five strategy-proof mechanisms examined in Section 3, the pivotal mechanism, the serial cost sharing mechanism, the second price auction, the Condorcet voting scheme, and the uniform rule fail to satisfy the rectangular property, so that they have a continuum of Nash equilibria as illustrated in Figures 1-6.

In the early literature on implementation, it was pointed out that even if an SCF *f* is implementable in dominant strategies, it may not be implemented by its associated direct revelation mechanism: it may be necessary to use more complicated "indirect" mechanisms (Dasgupta, Hammond, and Maskin (1979), Repullo (1985)). However, Repullo (1985) showed that if the direct revelation mechanism associated with the SCF *f* does not implement *f* in dominant strategies, then *f* is not Nash implementable by any mechanism. Hence, it cannot be securely implemented. Conversely, suppose the SCF *f* is securely implemented by some mechanism. Then by Proposition 1 and lemma 1, *f* satisfies strategy-proofness and the rectangular property. Hence, by lemma 2, *f* is securely implemented by its associated direct revelation mechanism. Thus, we have a *revelation principle for secure implementation:*

Theorem 2. *An SCF is securely implementable if and only if it is securely implemented by its associated direct revelation mechanism.*

The implication of this revelation principle is that we can limit our attention to the set of direct mechanisms. Direct mechanisms are somewhat natural and easy to explain to experimental subjects, which may add to their appeal.

5. Applications to Public Good Economies with Quasi-Linear Preferences: the Class of Groves-Clarke Mechanisms

In this section, we apply our characterization result, Theorem 1, to public good economies with quasi-linear preferences. We find that none of strategy-proof and efficient SCF's is securely implementable if public goods are discrete. On the other hand, strategy-proof and efficient SCF's are securely implementable by Grove-Clarke mechanisms with single-peaked preferences.

Let the set of alternatives be

$$A = \{(y, t_1, \ldots, t_n) | y \in Y, t_i \in \Re, \forall i\},$$

where $Y \subseteq \Re$ is a production possibility set, $y \in Y$ is an output level of a public good, and t_i is a transfer of a private good to agent i. For simplicity, we assume that there is no cost involved in producing y. Each agent has selfish and quasi-linear preferences:

$$u_i(y, t_1, \ldots, t_n) = u_i(y, t_i) = v_i(y) + t_i, \quad i \in I.$$

The class of valuation functions admissible for agent i is denoted by V_i. Let $V \equiv \times_{i \in I} V_i$.

Consider an SCF f satisfying the efficiency condition on the public good provision:

(5.1) $y^f(v_1, v_2, \ldots, v_n) \in \arg\max_{y \in Y} \sum_{i \in I} v_i(y)$ for all $v \in V$,

where $y^f(v)$ denotes the level of the public good recommended by the SCF f for the

profile v.

The following result is well-known:

Proposition 2 (*Clarke (1971), Groves (1973), Green and Laffont (1979)*). *An SCF f satisfying* (5.1) *is dominant strategy implementable if and only if f satisfies*

(5.2) $t_i^f(v_1, v_2, \dots, v_n) = \sum_{j \neq i} v_j(y^f(v_1, v_2, \dots, v_n)) + h_i(v_1, \dots, v_{i-1}, v_{i+1} \dots, v_n) \quad \forall v \in V, \forall i$,

where $t_i^f(v)$ *represents the transfer to agent i recommended by the SCF f for the utility profile v and* h_i *is some arbitrary function which does not depend on* v_i.

A direct revelation mechanism satisfying (5.1) and (5.2) is called a *Groves-Clarke mechanism*. Proposition 2 says that we can focus on the class of Groves-Clarke mechanisms for implementation of an efficient SCF in dominant strategy equilibria.

We now check whether or not secure implementation of an efficient SCF is possible on two different environments.

1) *Finite Public Projects*

Suppose that Y is a finite set, called a set of public project choices, and $0 \in Y$, where 0 means that no public project is produced and $y \in Y$ such that $y \neq 0$ means that some public project is produced. Let the class of valuation functions admissible for agent i be $V_i = \{v_i : Y \to \Re \mid v_i(0) = 0, v_i(y) \in \Re, \forall y \in Y\}$ for $i \in I$.

Lemma 3. Let Y be a finite set of public project choices with $0 \in Y$, *and* $V_i = \{v_i : Y \to \Re \mid v_i(0) = 0, v_i(y) \in \Re, \forall y \in Y\}$ *for each* $i \in I$. *Then any SCF satisfying (5.1) and (5.2) fails to satisfy the rectangular property.*

Proof of Lemma 3: Take some $b \in Y$ with $b \neq 0$. Take a valuation profile $v \in V$ such that

$$\max_{y \in Y - \{0\}} \sum_{i \in I} v_i(y) = \sum_{i \in I} v_i(b) < 0 \text{ and } \max_{y \in Y - \{0\}} \sum_{j \neq i} v_j(y) = \sum_{j \neq i} v_j(b) < 0 \quad \forall i \in I.$$

15

Given such a profile $v \in V$, construct another profile $\tilde{v} \in V$ such that

$$\max_{y \in Y-\{0\}} \tilde{v}_i(y) = \tilde{v}_i(b) = \left| \sum_{j \neq i} v_j(b) \right| / n \quad \forall i \in I .$$

Since $\max_{y \in Y-\{0\}} \sum_{i \in I} v_i(y) < \sum_{i \in I} v_i(0) = 0$,

(5.3) $y^f(v) = 0$ and

$$t_i^f(v) = \sum_{j \neq i} v_j(y^f(v)) + h_i(v_{-i}) = h_i(v_{-i}) \text{ for all } i \in I .$$

Also notice that

$$\max_{y \in Y-\{0\}} [\tilde{v}_i(y) + \sum_{j \neq i} v_j(y)] \leq \max_{y \in Y-\{0\}} \tilde{v}_i(y) + \max_{y \in Y-\{0\}} \sum_{j \neq i} v_j(y)$$

$$= \left| \sum_{j \neq i} v_j(b) \right| / n + \sum_{j \neq i} v_j(b)$$

$$< 0.$$

Hence,

(5.4) $y^f(\tilde{v}_i, v_{-i}) = 0$ and

$$t_i^f(\tilde{v}_i, v_{-i}) = \sum_{j \neq i} v_j(y^f(\tilde{v}_i, v_{-i})) + h_i(v_{-i}) = h_i(v_{-i}) \text{ for all } i \in I .$$

By (5.3) and (5.4),

$$\tilde{v}_i(y^f(v)) + t_i^f(v) = \tilde{v}_i(y^f(\tilde{v}_i, v_{-i})) + t_i^f(\tilde{v}_i, v_{-i}) \text{ for all } i \in I .$$

However, since $\max_{y \in Y-\{0\}} \sum_{i \in I} \tilde{v}_i(y) = \sum_{i \in I} \tilde{v}_i(b) = \left| \sum_{j \neq i} v_j(b) \right|$, $y^f(\tilde{v}) = b \neq y^f(v)$. Q.E.D.

Theorem 1, Proposition 2, and Lemma 3 together imply the following

impossibility result:

Theorem 3. *Let Y be a finite set of public project choices with $0 \in Y$, and*

$V_i = \{v_i : Y \to \Re \,|\, v_i(0) = 0, v_i(y) \in \Re, \forall y \in Y\}$ *for each $i \in I$. If an efficient SCF satisfying (5.1)*

is dominant strategy implementable, then it is not securely implementable.

In other words, for any mechanism implementing an efficient SCF in dominant

strategy equilibria, the set of Nash equilibrium outcomes is strictly larger than that of

dominant strategy equilibrium outcomes if the number of public project choices is finite.

2) *Continuum Public Projects with Single-Peaked Preferences*

Suppose that $Y = \Re$ and for all $i \in I$,

$$V_i = \{v_i : \Re \to \Re | v_i(y) = -(y - r_i)^2, r_i \in \Re\},$$

where r_i is agent i's most preferred level of the public good. We can represent these single-peaked preferences by the r_i instead of the v_i. The optimal output level of the public good satisfying (5.1) is given by

$$y^f(r_1, r_2, \ldots, r_n) = \frac{1}{n} \sum_{i \in I} r_i.$$

We will prove that any SCF f meeting (5.1) and (5.2) satisfies the rectangular property. Take two profiles $r, \tilde{r} \in \Re^n$. Suppose that

$$\tilde{v}_i(y^f(r)) + t_i^f(r) = \tilde{v}_i(y^f(\tilde{r}_i, r_{-i})) + t_i^f(\tilde{r}_i, r_{-i}) \text{ for all } i \in I.$$

We can rewrite these equations as

$$(A - \tilde{r}_i)^2 + \sum_{j \neq i}(A - r_j)^2 = (A + B_i - \tilde{r}_i)^2 + \sum_{j \neq i}(A + B_i - r_j)^2 \text{ for all } i \in I,$$

where $A = \sum_{\ell=1}^{n} r_\ell / n$ and $B_i = (\tilde{r}_i - r_i) / n$ for all $i \in I$. By using the above equations, it is not hard to see that $r_i = \tilde{r}_i$ for all $i \in I$. Therefore, $y^f(r) = y^f(\tilde{r}_i, r_{-i})$ and $t_i^f(r) = t_i^f(\tilde{r}_i, r_{-i})$ for all $i \in I$. Accordingly, the rectangular property is satisfied.

This result together with Theorem 1 and Proposition 2 imply the following corollary.

Corollary 1. *Suppose that $Y = \Re$ and for all $i \in I$, $V_i = \{v_i : \Re \to \Re | v_i(y) = -(y - r_i)^2, r_i \in \Re\}$. Then any SCF satisfying (5.1) and (5.2) is securely implementable.*

6. Dominant Strategy Implementation versus Secure Implementation

By applying the revelation principle for secure implementation, Theorem 2, we can clarify the difference between dominant strategy implementability and secure

implementability. The literature contains examples of social choice functions that are dominant strategy implemented by some mechanism g, but not dominant strategy implemented by the associated direct revelation mechanism (e.g., Repullo (1985)). By Theorem 2, each of these examples in the literature gives an example of an SCF that can be dominant strategy implemented, but cannot be securely implemented. This illustrates the significant difference between the set of dominant strategy implementable social choice functions and the set of securely implementable social choice functions. In fact, Theorem 2 implies that the set of securely implementable SCF's is a subset of the set of SCF's that are dominant strategy implementable *by their associated direct revelation mechanisms*. This latter set is, of course, a strict subset of the set of dominant strategy implementable SCF's, which in turn is a strict subset of the set of strategy proof SCF's. To better understand the difference between the set of securely implementable SCF's and the set of SCF's that are dominant strategy implementable by their associated direct revelation mechanisms, we introduce the following definition.

Definition 4. The SCF f satisfies *preference sensitivity* if for all $u, u' \in U$ and all $i \in I$, if $f(u_i, u_{-i}) \neq f(u'_i, u_{-i})$, then there is some u''_{-i} such that $u_i(f(u_i, u''_{-i})) > u_i(f(u'_i, u''_{-i}))$.

Preference sensitivity together with strategy-proofness are necessary and sufficient for dominant strategy implementability by the direct revelation mechanism:

Theorem 4. *An SCF is dominant strategy implemented by its associated direct revelation mechanism if and only if it satisfies strategy-proofness and preference sensitivity.*

Proof of Theorem 4. Suppose the SCF f satisfies strategy-proofness and preference sensitivity. Consider the associated direct revelation mechanism. Suppose agent i's true preference is u_i. By strategy proofness, it is dominant to announce the truth u_i.

Suppose announcing a different preference u_i' is another dominant strategy. If $f(u_i, u_{-i}) \neq f(u_i', u_{-i})$ for some u_{-i}, then by preference sensitivity there is u_{-i}'' such that $u_i(f(u_i, u_{-i}'')) > u_i(f(u_i', u_{-i}''))$. Therefore, announcing u_i' is in fact dominated by announcing u_i, which is a contradiction. Hence, $f(u_i, u_{-i}) = f(u_i', u_{-i})$ for all u_{-i} after all, so agent i's lie (i.e. to say u_i') cannot ever affect the outcome. Hence, f is dominant strategy implemented.

Suppose the SCF f is dominant strategy implemented by its associated direct revelation mechanism. By Proposition 1, f is strategy-proof. It remains to show f satisfies preference sensitivity. Take any $u, u' \in U$ and $i \in I$. Suppose $f(u_i, u_{-i}) \neq f(u_i', u_{-i})$. Then announcing u_i' is dominated by announcing u_i when agent i's true preference is u_i, so that there is u_{-i}'' such that $u_i(f(u_i, u_{-i}'')) > u_i(f(u_i', u_{-i}''))$.   Q.E.D.

Corollary 2. *If an SCF fails to satisfy preference sensitivity, then it is not securely implementable by any mechanism.*

Notice that preference sensitivity does not imply that each player will have a unique dominant strategy in the revelation mechanism. Repullo (1985) gave an example of a dominant strategy implementable SCF that is not implemented by its associated revelation mechanism, and hence violates preference sensitivity by Theorem 4. It is easy to check that all five SCF's discussed in Section 3 do satisfy preference sensitivity. Since all five violate the rectangular property, the difference between the two conditions is highly significant.

7. Concluding Remarks

Many researchers believe that if truth telling is a dominant strategy, then every agent will adopt it. However, we believe this issue should be decided by experiments. In Cason, Saijo, Sjöström and Yamato (2003), we conducted experiments on two

strategy-proof mechanisms: the pivotal mechanism with a binary public project that has a continuum of Nash equilibria, and a Groves mechanism with single-peaked preferences that has a unique Nash equilibrium. We found a clear difference on the choice of dominant strategies between the two: the frequency of dominant strategy equilibria chosen by subjects was 50% in the pivotal mechanism experiment, while it was 81% in the Groves mechanism experiment.[6] We are currently conducting further experiments.

---

[6] In both experiments, we used payoff tables only. No details of the rules of mechanisms were given to subjects. In other words, we avoided possible complexity resulting from non-understanding of the rules of mechanisms by subjects.

References

Attiyeh, G., Franciosi, R. and Isaac, R. M. (2000) "Experiments with the Pivotal Process for Providing Public Goods," *Public Choice*, 102, 95-114.

Barberà, S. and Jackson, M. O. (1995) "Strategy-proof Exchange," *Econometrica*, 63, 51-87.

Barberà, S., Massó, J. and Serizawa, S. (1998) "Strategy-proof Voting on Compact Ranges," *Games and Economic Behavior*, 25, 272-91.

Border, K. and Jordan, J. (1983) "Straightforward Elections, Unanimity and Phantom Voters," *Review of Economic Studies*, 50, 153-70.

Cason, T., Saijo, T., Sjöström , T, and Yamato, T. (2003) "Secure Implementation Experiments: Do Strategy-Proof Mechanisms Really Work?" Social Science Working Paper No. 1165, California Institute of Technology.

Clarke, E. H. (1971). "Multipart Pricing of Public Goods," *Public Choice*, 2, 19-33.

Gibbard, A. (1973). "Manipulation of Voting Schemes," *Econometrica*, 41, 587-601.

Green, J. R. and Laffont, J.-J. (1979). *Incentives in Public Decision Making.* Amsterdam: North- Holland.

Groves, T. (1973). "Incentives in Teams," *Econometrica*, 41, 617-31.

Kagel, J. H., Harstad, R. M. and Levin, D. (1987) "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study," *Econometrica,* 55, 1275-1304.

Kagel, J. H. and Levin, D. (1993) "Independent Private Value Auctions: Bidder Behavior in First- Second- and Third-Price Auctions with Varying Number of Bidders," *Economic Journal*, 103, 868-879.

Kawagoe, T. and Mori, T. (2001) "Can the Pivotal Mechanism Induce Truth-Telling? An Experimental Study," *Public Choice*, 108, 331-354.

Maskin, E. (1979) "Incentive Schemes Immune to Group Manipulation," mimeo.

Moulin, H. (1980) "On Strategy-proofness and Single Peakedness," *Public Choice*, 35, 437-455.

Moulin, H. (1994) "Serial Cost-Sharing of Excludable Public Goods," *Review of Economic Studies*, 61, 305-325.

Repullo, R. (1985) "Implementation in Dominant Strategies under Complete and Incomplete Information," *Review of Economic Studies*, 52, 223-229.

Schummer, J. and Vohra, R. V. (2001) "Strategy-proof Location on a Network," *Journal of Economic Theory*, forthcoming.

Sprumont. Y. "The Division Problem with Single-peaked Preferences: A Characterization of the Uniform Allocation Rule," *Econometrica*, 59, 509-519.

Vickrey, W. (1961). "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, 16, 8-37.

Yamato, T. (1993) "Double Implementation in Nash and Undominated Nash Equilibria," *Journal of Economic Theory*, 59, 311-323.
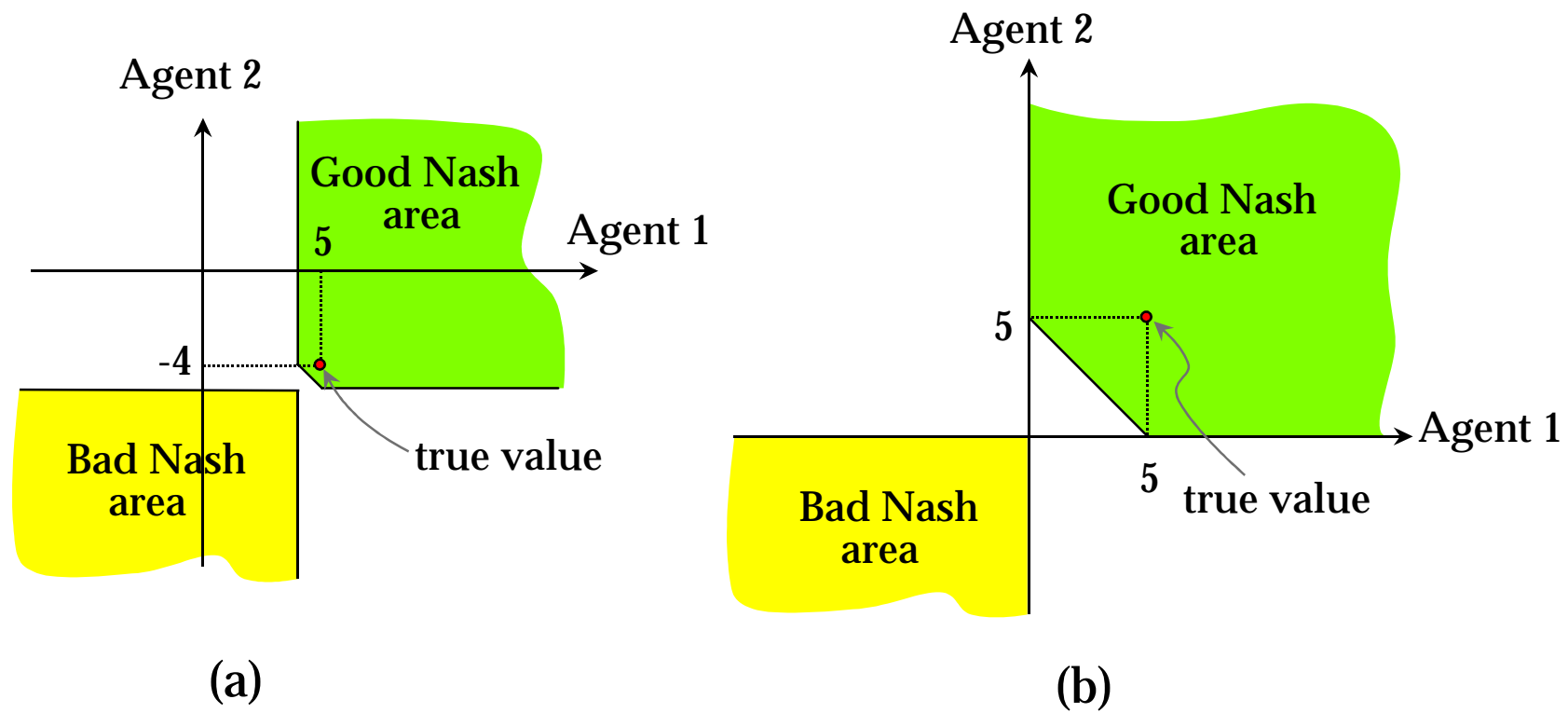
Figure 1: Equilibria of the Pivotal Mechanism
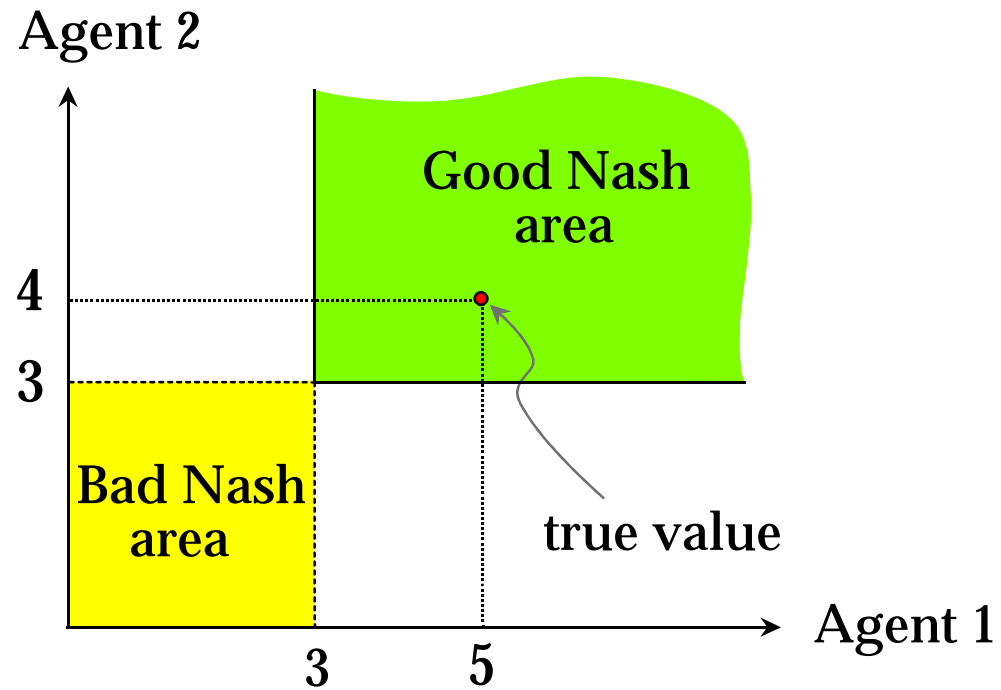
Figure 2: Equilibria of the Serial Cost Sharing Mechanism
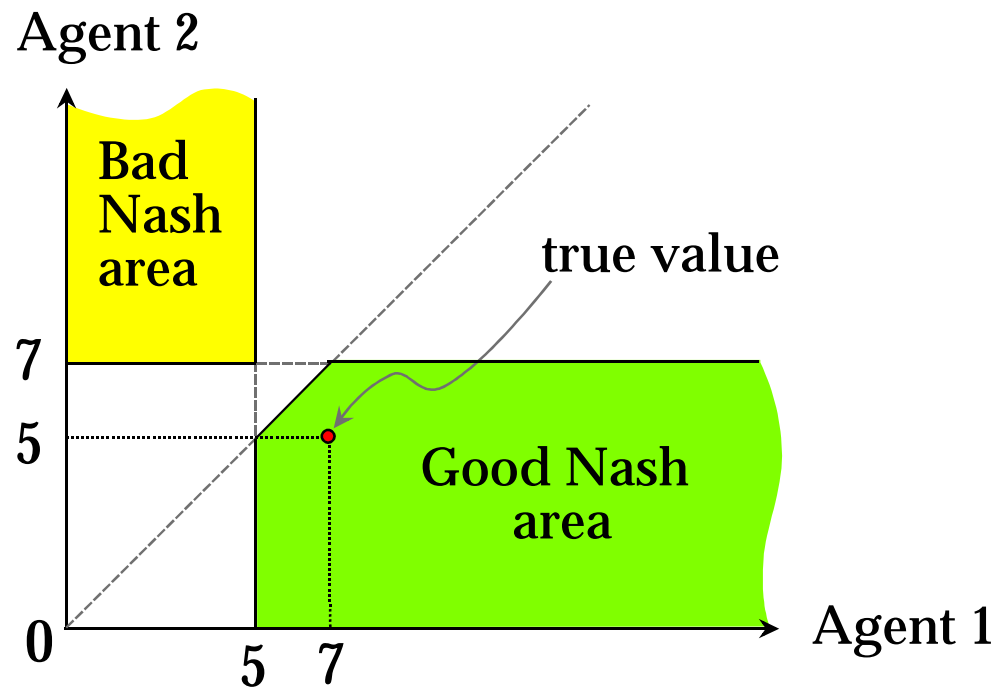
Figure 3: Equilibria of the Second Price Auction
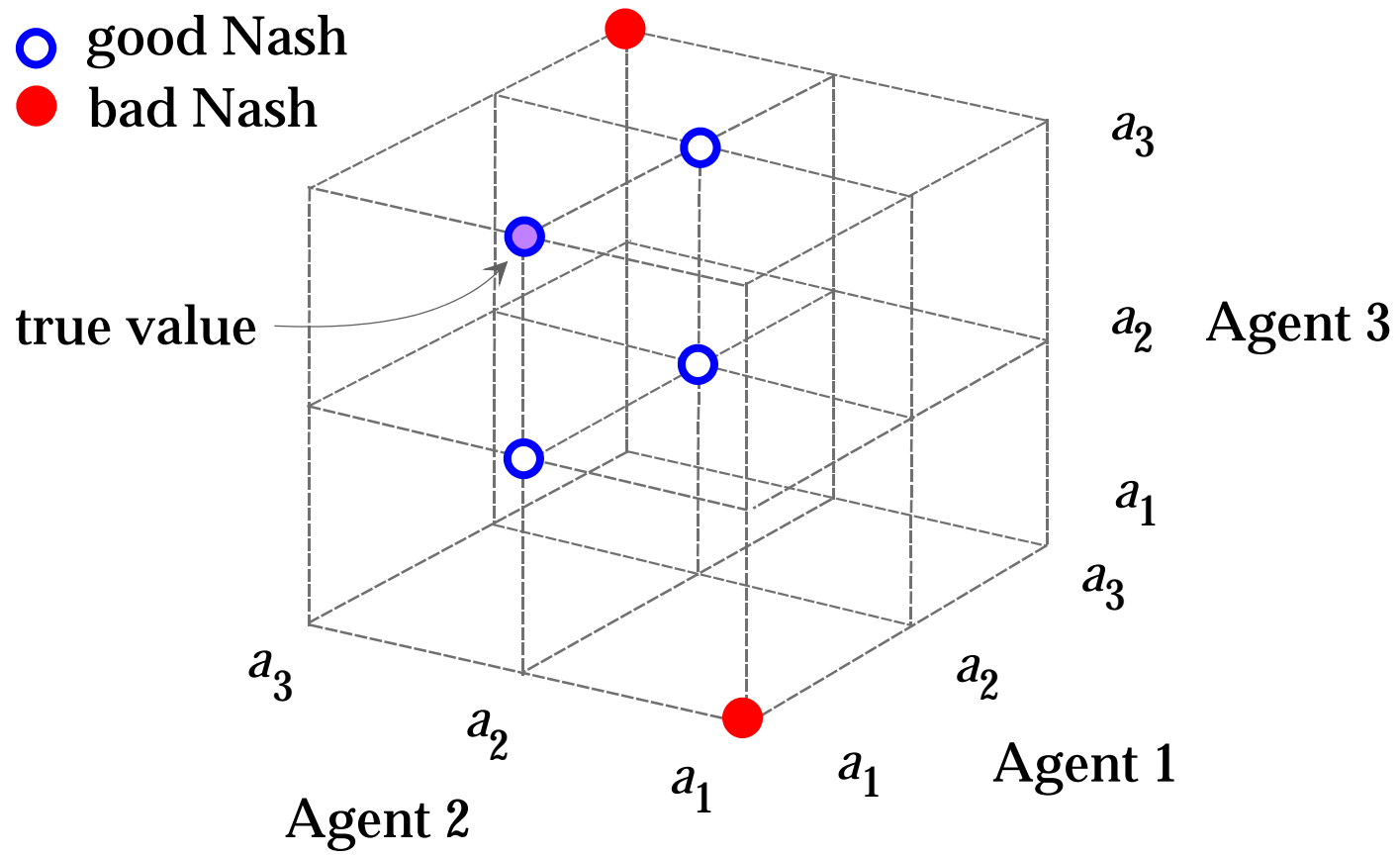
Figure 4: Equilibria of the Condorcet Voting Scheme
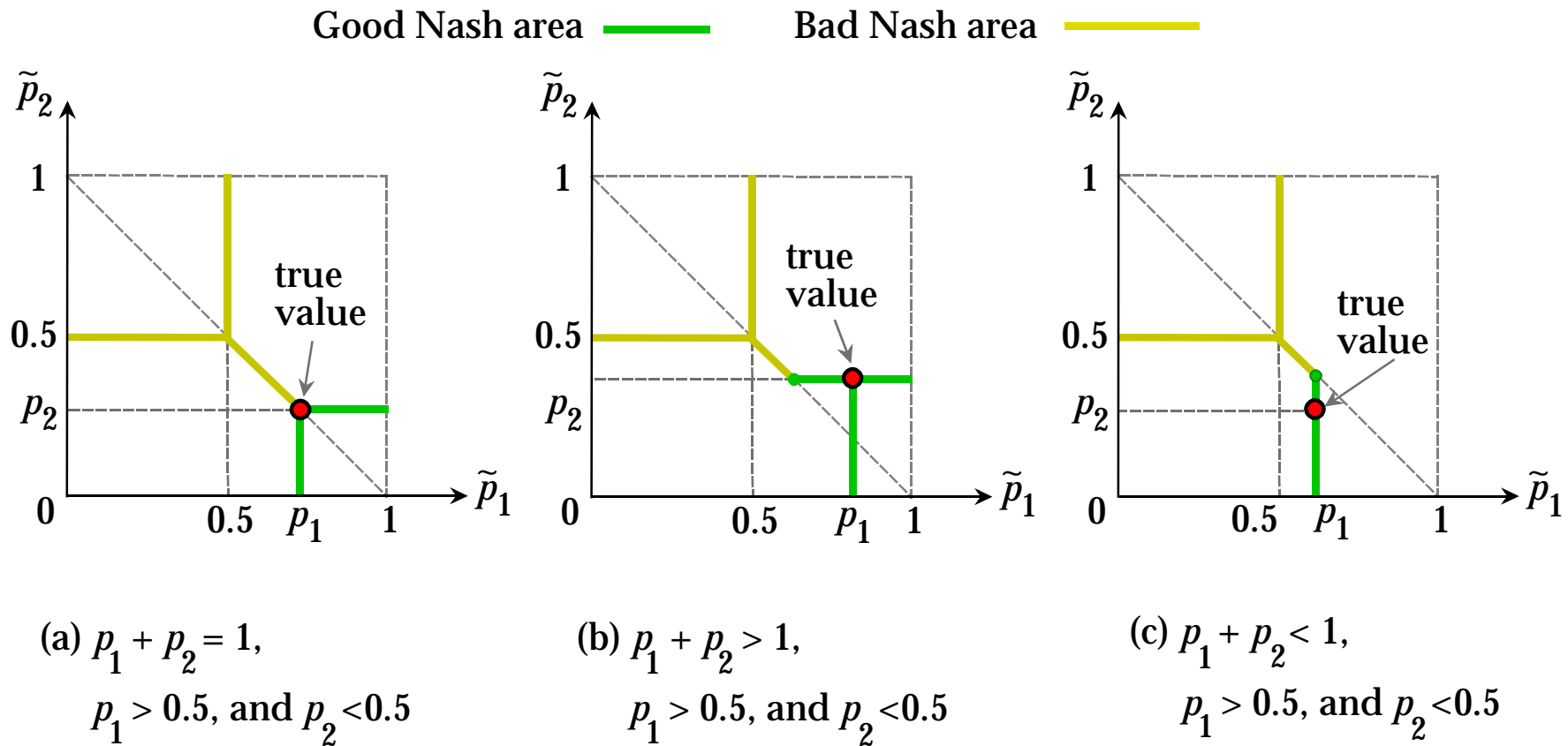
Figure 5: Equilibria of the Uniform Rule (1)
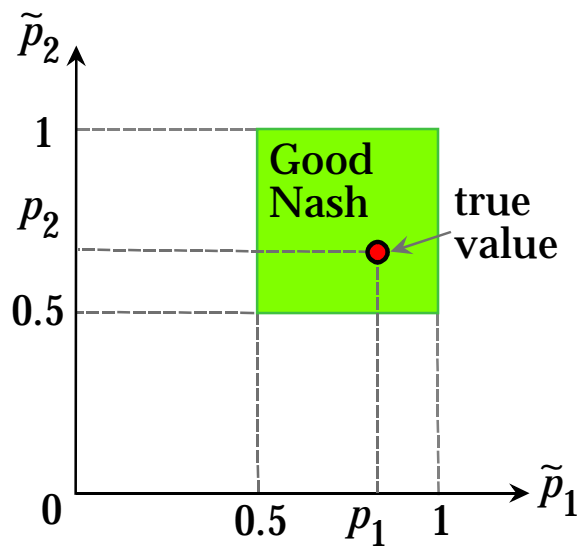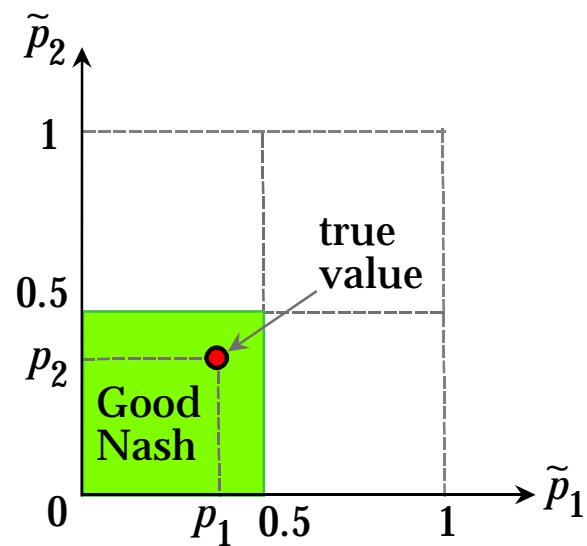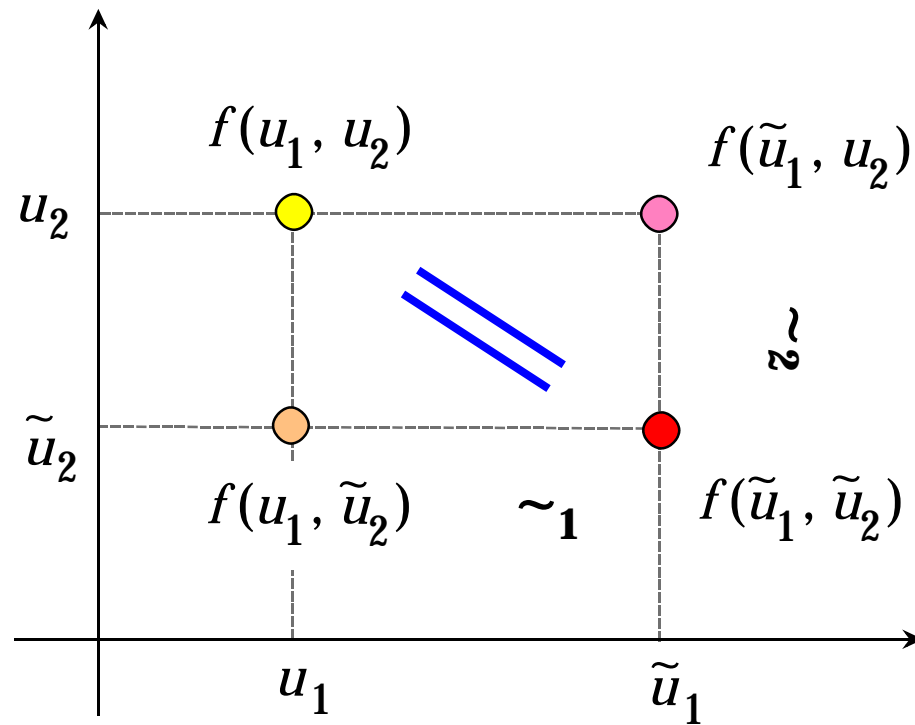
Figure 6: Equilibria of the Uniform Rule (2)

Figure 7: Rectangular Property