Secure Implementation[†]

Tatsuyoshi Saijo*

Tomas Sjöström**

and

Takehiko Yamato***

June, 1998/Revised July 22, 2005

[†] Research was partially supported by the Grant in Aid for Scientific Research 15310023 of the Ministry of Education, Science and Culture in Japan, the Japan Securities Scholarship Foundation, and the Abe Fellowship. We thank Salvador Barberà, Dirk Bergemann, Ken Binmore, Hervé Moulin, Shinji Ohseto, Thomas Palfrey, and Jun Wako for useful comments.

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

** Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, New Jersey 08901-1248, USA. E-mail: tsjostrom@economics.rutgers.edu

*** Department of Value and Decision Science and Department of Social Engineering, 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, this concept has serious drawbacks. In particular, many strategy-proof mechanisms have multiple Nash equilibria, some of which produce the wrong outcome. A possible solution to this problem is to require double implementation in Nash equilibrium and in dominant strategies, i.e., secure implementation. We characterize securely implementable social choice functions, and compare our results with dominant strategy implementation. In standard quasi-linear environments with divisible private or public goods, there exist Pareto efficient (non-dictatorial) social choice functions that can be securely implemented. But in the absence of side-payments, secure implementation is incompatible with Pareto efficiency.

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

Correspondent:

Tomas Sjöström
Department of Economics
Rutgers University
75 Hamilton Street, New Brunswick
New Jersey 08901-1248, USA.
E-mail: tsjostrom@economics.rutgers.edu

1. Introduction

Strategy-proofness, requiring that truth-telling is a dominant strategy, is a standard concept in social choice theory. Although it seems natural that an agent will tell the truth if it is a dominant strategy to do so, there are some problems. 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 multiple Nash equilibria, some of which produce the wrong outcome. Third, experimental evidence shows that some 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. This may require the use of indirect mechanisms. However, Repullo (1985) showed that if a 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. This leads to the second problem: mechanisms for dominant strategy implementation may have "bad" Nash equilibria. For this reason, Repullo (1985) suggested that the concept of dominant strategy implementation should be replaced by Nash or Bayesian Nash implementation. We agree that the existence of "bad" (Bayesian) Nash equilibria is problematic. However, 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 a failure to coordinate on a (Bayesian) Nash equilibrium. Moreover, a problematic aspect of Bayesian Nash implementation is that it typically requires the mechanism designer to know the common prior of the players.

It seems clear that the standard concepts – dominant strategy implementation and (Bayesian) Nash implementation – cannot provide 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 dual advantages. First, with dominant strategies, strategic uncertainty is not important. Second, the mechanism implements the social choice rule in Bayesian Nash equilibria, for any common prior the players may hold. There is no possibility of getting stuck at a "bad" equilibrium.

A social choice function is *securely implementable* if there exists a game form 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*. We show that many quasi-linear economic environments with continuous private or public goods admit securely implementable non-dictatorial social choice functions that maximize social surplus. However, in a standard single-peaked voting model without side-payments, any securely implementable social choice rule must be either dictatorial or Pareto inefficient. This negative result holds even for multi-valued social choice correspondences. In a quasi-linear environment with a *discrete* social decision, such as whether or not to implement an indivisible public project, some interesting non-dictatorial social choice correspondences can be securely implemented, but none of them maximizes the social surplus.

Our hope is that secure implementation may lead to some progress on the third problem mentioned above, the rather negative experimental evidence. 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 question of whether secure mechanisms work well in experiments is investigated in a companion paper (Cason, Saijo, Sjöström and Yamato (2003)).

The remainder of the paper is organized as follows. We give notation and definitions in Section 2. We characterize secure implementability in Section 3. In Section 4 we discuss the relationship between non-bossiness, dominant strategy implementation and secure implementation. In Section 5, we consider "robust" Bayesian Nash implementation. In Section 6, we show the possibility of secure implementation in economies with quasi-linear preferences and divisible public and private goods. Sections 7 and 8 discuss the difficulty of secure implementation with discrete social decisions, and in the absence of side-payments. Concluding remarks are in Section 9.

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 $L(a, u_i) \equiv \{b \in A \mid u_i(a) \geq u_i(b)\}$ 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 $N^g_A(u)$ be the set of Nash equilibrium outcomes of g at u, i.e., $N^g_A(u) \equiv \{a \in A \mid$ 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 $u \in U$, $f(u) = N^g_A(u)$. f is *Nash implementable* if there exists a mechanism which implements f in Nash equilibria. 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 $u \in U$. We will sometimes abuse terminology by not distinguishing between the SCF f and the direct revelation mechanism associated with f.

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 $DS^g_i(u_i)$ 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 $i \in I$, $s_i \in DS^g_i(u_i)$. Let $DS^g(u)$ be the set of dominant strategy equilibria of g at u. Also, let $DS^g_A(u)$ be the set of dominant strategy equilibrium outcomes of g at u, i.e., $DS^g_A(u) \equiv \{a \in A \mid$ 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 $u \in U$, $f(u) = DS^g_A(u)$. 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 $\tilde{u}_{-i} \in U_{-i}$, $u_i(f(u_i, \tilde{u}_{-i})) \geq u_i(f(\tilde{u}_i, \tilde{u}_{-i}))$. 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. Secure Implementation: A Characterization and a Revelation Principle

We introduce the following new concept of implementation.

Definition 1. The mechanism *g securely implements the SCF f* if for all $u \in U$, $f(u) = DS_A^g(u) = N_A^g(u)$.[1] 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 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 1 in which $n = 2$ and (u_1, u_2) is the true preference profile. Suppose

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

---

[1] 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.

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 (3.1) that

$u_1(f(\tilde{u}_1,\tilde{u}_2)) = u_1(f(u_1,\tilde{u}_2)) \geq u_1(f(u_1',\tilde{u}_2))$ for all $u_1' \in U_1$.

That is, reporting $\tilde{u}_1$ is one of agent 1's best responses at u_1 when agent 2 reports $\tilde{u}_2$.

Next suppose that

(3.2) $\quad u_2(f(\tilde{u}_1,u_2)) = u_2(f(\tilde{u}_1,\tilde{u}_2))$.

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

$u_2(f(\tilde{u}_1,\tilde{u}_2)) \geq u_2(f(\tilde{u}_1,u_2'))$ for all $u_2' \in U_1$, that is, reporting $\tilde{u}_2$ is one of agent 2's best responses when agent 1 reports $\tilde{u}_1$. Therefore, $f(\tilde{u}_1,\tilde{u}_2)$ 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 $f(u_1,u_2) = f(\tilde{u}_1,\tilde{u}_2)$ if (3.1) and (3.2) holds.

Link to Figure 1

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

Definition 2. The SCF *f* satisfies the *rectangular property* if for all $u,\tilde{u} \in U$, if $u_i(f(\tilde{u}_i,\tilde{u}_{-i})) = u_i(f(u_i,\tilde{u}_{-i}))$ for all $i \in I$, then $f(\tilde{u}) = f(u)$.

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

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

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

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

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

(3.4) $\quad g(s_1(\tilde{u}_1), \ldots, s_n(\tilde{u}_n)) = f(\tilde{u})$.

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 $(s_i(u_i), s_{-i}(\tilde{u}_{-i})) \in DS^g(u_i, \tilde{u}_{-i})$, where $s_{-i}(\tilde{u}_{-i}) = (s_j(\tilde{u}_j))_{j \neq i}$. By dominant implementability,

(3.5) $\quad g(s_i(u_i), s_{-i}(\tilde{u}_{-i})) = f(u_i, \tilde{u}_{-i})$.

By (3.3), (3.4), and (3.5),

(3.6) $\quad u_i(g(s_i(u_i), s_{-i}(\tilde{u}_{-i}))) = u_i(g(s_1(\tilde{u}_1), \ldots, s_n(\tilde{u}_n)))$.

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

(3.7) $\quad g(S_i, s_{-i}(\tilde{u}_{-i})) \subseteq L(g(s_i(u_i), s_{-i}(\tilde{u}_{-i})), u_i)$.

By (3.6) and (3.7), $g(S_i, s_{-i}(\tilde{u}_{-i})) \subseteq L(g(s_i(\tilde{u}_i), s_{-i}(\tilde{u}_{-i})), u_i)$. Since this holds for any $i \in I$, $(s_1(\tilde{u}_1), \ldots, s_n(\tilde{u}_n)) \in N^g(u)$. By Nash implementability and (3.4), $f(u) = g(s_1(\tilde{u}_1), \ldots, s_n(\tilde{u}_n)) = f(\tilde{u})$. Q.E.D.

Next we show that strategy-proofness and the rectangular property are not only necessary, but also sufficient for secure implementability.

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: By strategy-proofness, for all $u \in U$, $f(u) \in DS_A^g(u)$. We will prove that for all $u \in U$, $N_A^g(u) = f(u)$. Since $\varnothing \neq DS_A^g(u) \subseteq N_A^g(u)$, that suffices to prove the lemma.

Let $u \in U$ be given. Take any $s = \tilde{u} \in N^g(u)$. We show that $g(s) = f(u)$, i.e., $f(\tilde{u}) = f(u)$. Since $\tilde{u} \in N^g(u)$, (3.8) $u_i\,(f(\tilde{u}_i, \tilde{u}_{-i})) \geq u_i\,(f(u_i, \tilde{u}_{-i}))$ for all $i \in I$.

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

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

By (3.8) and (3.9), $u_i\,(f\,(\tilde{u}_i, \tilde{u}_{-i})) = u_i\,(f\,(u_i, \tilde{u}_{-i}))$ for all $i \in I$. By the rectangular property, $f(\tilde{u}) = f(u)$. Q.E.D.

By Proposition 1, Lemmas 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.*

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, 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.

4. Non-Bossiness, Dominant Strategy Implementation and Secure Implementation

To further study the set of securely implementable social choice functions, we need the idea of *non-bossiness*. Intuitively, non-bossiness implies that no one can change the outcome without changing her own utility. Satterthwaite and Sonnenschein (1981) first introduced a definition of non-bossiness for economic environments.[2] For general environments, consider the following definitions.

Definition 3. The SCF *f* satisfies *non-bossiness* 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 $u_i(f(u_i, u_{-i})) \neq u_i(f(u_i', u_{-i}))$.

Definition 4. The SCF *f* satisfies *weak non-bossiness* 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 $u_i(f(u_i, u''_{-i})) \neq u_i(f(u_i', u''_{-i}))$.

The rectangular property is stronger than non-bossiness.

Proposition 2. *If an SCF satisfies the rectangular property, then it satisfies non-bossiness.*

Proof: Suppose the SCF *f* satisfies the rectangular property, and $u_j(f(u_j', u_{-j})) = u_j(f(u_j, u_{-j}))$ for some *j*. Let u'' be such that $u'' = (u_j', u_{-j})$. We need to show $f(u'') = f(u)$. Now $u_j(f(u'')) = u_j(f(u_j', u_{-j})) = u_j(f(u_j, u_{-j})) = u_j(f(u_j, u''_{-j}))$, and $(u_i'', u''_{-i}) = (u_i, u''_{-i})$ for all $i \neq j$. So we have $u_i(f(u_i'', u''_{-i})) = u_i(f(u_i, u''_{-i}))$ for all $i \in I$. By the rectangular property, $f(u'') = f(u)$. Q.E.D.

---

[2] Our definition of non-bossiness is slightly stronger than Satterthwaite-Sonnenschein's original condition, when applied to economic environments. Satterthwaite and Sonnenschein's original definition was that the SCF *f* satisfies non-bossiness 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 $f_i(u_i, u_{-i}) \neq f_i(u_i', u_{-i})$, where $f_i(u)$ denotes the consumption bundle agent *i* receives at the allocation

Thus, any securely implementable SCF must be non-bossy. On the other hand, non-bossiness does not imply the rectangular property, even in combination with strategy-proofness (an example is provided in Section 6). However, it turns out that *weak* non-bossiness is enough to guarantee that a strategy-proof SCF can be *dominant strategy implemented*. Non-bossiness is a stronger condition than weak non-bossiness, so secure implementation is more difficult to achieve than dominant strategy implementation. For example, the Vickrey auction discussed in Section 7 satisfies weak non-bossiness, but violates non-bossiness. (Notice that in general, weak non-bossiness does not imply that each player will have a unique dominant strategy in the revelation mechanism.)

Theorem 3. *An SCF is dominant strategy implemented by its associated direct revelation mechanism if and only if it satisfies strategy-proofness and weak non-bossiness.*

Proof: Suppose the SCF *f* satisfies strategy-proofness and weak non-bossiness. 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 weak non-bossiness 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*

$f(u) = (f_i(u))_{i \in I}$ recommended by the SCF *f* for the preference profile *u*. Mizukami and Wakayama (2004) discuss the importance of non-bossiness for dominant strategy implementation in exchange economies.

satisfies weak non-bossiness. 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.

5. Robust Bayesian Implementation without Knowledge of the Prior

In the standard theory of Bayesian mechanism design, the agents are assumed to have a common prior over the possible states of the world, and the planner is assumed to know this prior. For example, in Myerson's (1981) theory of optimal auctions, the optimal reserve price depends on the prior distribution. It has often been argued that the assumption that the planner knows the agents' common prior is too strong (e.g., Bergemann and Morris (2003, 2004)). It turns out that secure implementation guarantees that the planner does not need to know anything about the agents' beliefs. If strategy-proofness (i.e. dominant-strategy incentive compatibility) and the rectangular property hold, then if the agents play a Bayesian Nash equilibrium with any arbitrary prior whatsoever, the outcome will be socially optimal with probability one. Thus, the social planner can achieve implementation if the agents are Bayesian expected utility maximizers with a common prior, even if the planner is not sure about what the prior is.

Let U be a finite set of possible utility profiles, and let q be a common prior distribution over U. Fix a social choice function $f: U \to A$. The direct revelation mechanism associated with f, together with the common prior q, define a Bayesian game. A strategy for player i is a function $\sigma_i: U_i \to U_i$, with the following interpretation: when player i's true type is u_i, then he announces $\sigma_i(u_i)$. A strategy profile is a function $\sigma: U \to U$, with the following interpretation: when the true state is

u, the players announce $\sigma(u) = (\sigma_1(u_1), \sigma_2(u_2), \ldots, \sigma_n(u_n))$. Similarly, define $\sigma_{-i}(u_{-i})$ in the obvious way. We will allow an agent's beliefs about other agents' types to depend on the agent's own type. That is, after agent *i* learns his own preferences, he can update his prior belief using Bayes rule. If agent *i*'s true type is u_i, the the probability that he assigns to the other agents being u_{-i} is denoted $q_i(u_{-i}|u_i)$.

Bayesian Nash equilibrium is defined in the standard way. Notice that if the agents play a Bayesian Nash equilibrium σ in the direct revelation game, then if the true state is u, the mechanism will implement $f(\sigma(u))$. Given a prior distribution *q,* the direct revelation mechanism implements *f* in Bayesian Nash equilibria if the following two conditions hold: (i) a Bayesian Nash equilibrium σ exists, and (ii) for any Bayesian Nash equilibrium σ and any u such that $q(u) > 0$, $f(\sigma(u)) = f(u)$. Moreover, the direct revelation mechanism *robustly* implements *f* in Bayesian Nash equilibria if for any prior distribution *q,* the above two conditions hold. That is, robust implementation requires that the *same* mechanism works for *all q*.

We now prove that strategy-proofness and the rectangular property are necessary and sufficient for robust Bayesian Nash implementation by the direct revelation mechanism:

Theorem 4. *An SCF f is robustly implemented in Bayesian Nash equilibria by the direct revelation mechanism if and only if f satisfies strategy-proofness and the rectangular property.*

Proof: Suppose that *f* satisfies strategy-proofness and rectangular property. Since *f* is strategy-proof, the truthful strategy $\sigma*(u) = u$ is certainly a Bayesian Nash equilibrium. Now, suppose there is another Bayesian Nash equilibrium σ which is not truthful.

Take any $u*$ such that $\sigma(u*) = u' \neq u*$. We need to show that if $q(u*) > 0$, then $f(u') = f(u*)$. The Bayes-Nash property implies that

$$\sum_{u_{-i} \in U_{-i}} q_i(u_{-i}|u_i^*) u_i^*(f(\sigma_{-i}(u_{-i}), u_i')) \geq \sum_{u_{-i} \in U_{-i}} q_i(u_{-i}|u_i^*) u_i^*(f(\sigma_{-i}(u_{-i}), u_i^*))$$

for all i. By strategy-proofness, $u_i^*(f(\sigma_{-i}(u_{-i}), u_i')) \leq u_i^*(f(\sigma_{-i}(u_{-i}), u_i^*))$ for all u_{-i}. Therefore, we must in fact have

$$u_i^*(f(\sigma_{-i}(u_{-i}), u_i')) = u_i^*(f(\sigma_{-i}(u_{-i}), u_i^*))$$

for all u_{-i} such that $q_i(u_{-i}|u_i^*) > 0$. If each agent uses Bayes rule with the common prior q, then $q(u*) > 0$ implies $q_i(u_{-i}^*|u_i^*) > 0$ for all i. Therefore,

$$u_i^*(f(\sigma_{-i}(u_{-i}^*), u_i')) = u_i^*(f(\sigma_{-i}(u_{-i}^*), u_i^*))$$

for all i. But $\sigma(u*) = u'$, so

$$u_i^*(f(u'_{-i}, u_i')) = u_i^*(f(u'_{-i}, u_i^*)).$$

The rectangular property now implies $f(u') = f(u*)$.

Conversely, suppose that the direct revelation mechanism robustly implements *f* in Bayesian Nash equilibria. Then it follows from Proposition 1 and Lemma 1 that *f* is strategy-proof and satisfies the rectangular property: simply consider the special case of Bayesian Nash equilibrium when $q(u) = 1$ for some u, which corresponds to the case of complete information. Q.E.D.

By Theorems 2 and 4, we have the following equivalence result between secure implementability and robust Bayesian Nash implementability:

Corollary 1. *An SCF f is robustly implemented in Bayesian Nash equilibria by the direct revelation mechanism if and only if it is securely implemented by the direct revelation mechanism.*

6. Quasi-Linear Economic Environments

Let the set of alternatives be

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

where $y \in Y$ is a social decision, and t_i is a transfer to agent *i* of a private good called "money". The set of possible social decisions *Y* is a convex subset of $\Re^k$, for some *k*. (In the next section, we consider the case where *Y* is a discrete set). The cost of taking decision *y* (in terms of "money") is given by a differentiable and convex function *c*(*y*). Each agent $i \in I$ has quasi-linear preferences:

$$u_i(y, t_1, \dots, t_n) = v_i(y, \theta_i) + t_i .$$

Here v_i is a valuation function which is differentiable and concave in *y*, and θ_i is a real number representing agent *i*'s "type". For each *i*, the function v_i is given once and for all and only the type varies, so the preferences of the agents will be represented by the profile of types, $\theta = (\theta_1, \theta_2 \dots, \theta_n)$. The set of possible types for agent *i* is Θ_i. Let $\Theta \equiv \times_{i \in I} \Theta_i$. An SCF $f : \Theta \to A$ recommends, for each profile θ, a social decision $y^f(\theta)$ and a set of transfers. Let $t_i^f(\theta)$ denote the recommended transfer to agent *i*. Thus, $f(\theta) = (y^f(\theta), t_1^f(\theta), t_2^f(\theta) \dots, t_n^f(\theta))$. The *social surplus* is defined as

$$\sum_{i \in I} v_i(y, \theta_i) - c(y) \tag{6.1}$$

To avoid some technical issues, in this section we assume for all $\theta \in \Theta$, a unique *y* maximizes the social surplus (6.1). (This happens, for example, if each v_i is strictly concave in *y*). A direct revelation mechanism $f : \Theta \to A$ is a *Groves-Clarke mechanism* if for all $\theta \in \Theta$, $y^f(\theta)$ maximizes the social surplus, and the transfer function is given by, for all $i \in I$,

$$t_i^f(\theta) = \sum_{j \neq i} v_j(y^f(\theta), \theta_j) - c(y^f(\theta)) + \varphi_i(\theta_{-i}) \tag{6.2}$$

Here φ_i is some arbitrary function which does not depend on θ_i. It is well-known that Groves-Clarke mechanisms are strategy-proof (Clarke (1971), Groves (1973)). In many cases, for example if each v_i is differentiable in θ_i and each Θ_i is a convex set, any strategy-proof SCF that satisfies (6.1) must in fact also satisfy (6.2) (Holmström (1979)).

If the social surplus maximizing decision always occurs in the interior of Y (denoted int$Y$) then the rectangular property is equivalent to non-bossiness. Both properties reduce to the following: no agent should be able to change the profile of transfers without changing the social decision. This is shown in the following lemma.

Lemma 3. Suppose for all $\theta \in \Theta$, $y^f(\theta) \in \text{int}Y$ maximizes the social surplus. For any Groves-Clarke mechanism $f: \Theta \to A$, the following three conditions are equivalent: (i) f is non-bossy; (ii) for all $\theta, \theta' \in \Theta$ and $i \in I$, $f(\theta) = f(\theta_i', \theta_{-i})$ whenever $y^f(\theta) = y^f(\theta_i', \theta_{-i})$; (iii) f satisfies the rectangular property.

Proof: (i) implies (ii). Strategy proofness implies that if $y^f(\theta) = y^f(\theta_i', \theta_{-i})$, then $t_i^f(\theta) = t_i^f(\theta_i', \theta_{-i})$. Non-bossiness then implies $f(\theta) = f(\theta_i', \theta_{-i})$.

(ii) implies (iii). Suppose (ii) holds. Fix any profile θ, and let $y^* = y^f(\theta)$. Suppose $u_i(f(\theta_i', \theta_{-i})) = u_i(f(\theta))$ for all i. Since the surplus maximizing y is always unique, it is easy to see, using (6.2), that agent i desires $y = y^*$ uniquely. Therefore, if $u_i(f(\theta_i', \theta_{-i})) = u_i(f(\theta))$ then $y^f(\theta_i', \theta_{-i}) = y^*$. By property (ii), $f(\theta) = f(\theta_i', \theta_{-i})$. Since y^* is interior, and v is differentiable and concave in y, y^* can be found by solving the first order condition for maximizing (6.1). Since $y^f(\theta_i', \theta_{-i}) = y^f(\theta) = y^*$, we have $\partial v_i(y^*, \theta_i) / \partial y = \partial v_i(y^*, \theta_i') / \partial y$ for all i.

We know that $f(\theta_i', \theta_{-i}) = f(\theta)$ for all i. For $i = 1$, this yields $f(\theta_1', \theta_2, ..., \theta_n) = f(\theta)$. The first-order condition for maximizing (6.1) must be satisfied at y^* for profile $(\theta_1', \theta_2, ..., \theta_n)$. Since $\partial v_2(y^*, \theta_2) / \partial y = \partial v_2(y^*, \theta_2') / \partial y$, the first order condition is also satisfied at y^* for the profile $(\theta_1', \theta_2', \theta_3 ..., \theta_n)$. Thus,

$y^f(\theta_1', \theta_2', \theta_3 ..., \theta_n) = y^f(\theta_1', \theta_2, \theta_3 ..., \theta_n) = y*$. Property (ii) now implies $f(\theta_1', \theta_2', \theta_3 ..., \theta_n) = f(\theta_1', \theta_2, \theta_3 ..., \theta_n) = f(\theta)$. By sequentially replacing each θ_i by θ_i' in this manner, we find that $f(\theta') = f(\theta)$. Therefore, the rectangular property holds.

(iii) implies (i). This follows from Proposition 2. Q.E.D.

Example 1 shows that standard assumptions often guarantee non-bossiness.

Example 1: Production of a divisible public good.

The public good is one-dimensional, $Y = \Re_+$. There are two leading cases that have been studied in the literature. *Case 1:* $v_i(y, \theta_i) = \theta_i b(y)$, where b is a strictly concave function. To guarantee that the surplus maximizing y is strictly positive, suppose $b'(0) > 0$ and $c'(0) = 0$. *Case 2:* Let $g(x)$ be a function which is strictly concave, reaching a maximum at $x = 0$, and suppose $v_i(y, \theta_i) = g(y - \theta_i)$. There is no cost of producing the public good, $c(y) = 0$. This is the case of single-peaked preferences, where θ_i is agent i's "peak", i.e., his most preferred level of the public good. As long as all θ_i are strictly positive, the surplus maximizing level of the public good is strictly positive.

In both case 1 and case 2 of Example 1, if $y^f(\theta) = y^f(\theta_i', \theta_{-i})$ then $\theta_i' = \theta_i$, so obviously $f(\theta) = f(\theta_i', \theta_{-i})$. From Lemma 3 it follows that all Groves-Clarke mechanisms are non-bossy and will securely implement the social surplus maximizing decision.

Example 2 shows that corner solutions do not necessarily mean that secure implementation is impossible.

Example 2: Allocation of a divisible private good in fixed supply.

One unit of a divisible private good called "cake" is to be shared by the agents. (In addition, transfers of "money" are possible). The social decision is denoted

$y = (y_1, y_2, ..., y_n)$, where y_i is agent i's share of the cake. Feasibility requires $y \geq 0$ and $\sum_i y_i = 1$. Valuation functions are of the form $v_i(y, \theta_i) = \theta_i b(y_i)$, where b is a strictly increasing and strictly concave function, satisfying $b(0) = 0$. Suppose $\Theta_i = [\theta^{\min}, \theta^{\max}]$, where

(6.3) $\quad \theta^{\min} b'(0) > \theta^{\max} b'(1)$.

Inequality (6.3) guarantees that the social surplus is never maximized by giving all of the cake to one agent. However, with three or more agents, it may be optimal to give no cake to some agent, so Lemma 3 does not apply. The social surplus $\sum_i \theta_i b(y_i)$ is to be maximized subject to $y \geq 0$ and $\sum_i y_i = 1$. Let $\lambda > 0$ denote the Lagrange multiplier for the resource constraint. The maximum is found by solving the first order condition,

(6.4) $\quad \theta_i b'(y_i) \leq \lambda$, $y_i \geq 0$, $y_i(\lambda - \theta_i b'(y_i)) = 0$ for all i.

Suppose the function φ_i in (6.2) is a constant, so the transfer of money to agent i is

(6.5) $\quad t_i^f(\theta) = \sum_{j \neq i} \theta_j b(y_j^f(\theta)) + \text{constant}$.

We claim that in this case the Groves-Clarke mechanism satisfies the rectangular property. Indeed, suppose $u_i(f(\theta)) = u_i(f(\theta_i', \theta_{-i}))$ for all i. This implies that for all i, either $\theta_i' = \theta_i$ or agent i gets no cake, $y_i^f(\theta_i', \theta_{-i}) = y_i^f(\theta) = 0$. Therefore, the first order condition (6.4) still holds when θ is replaced by θ', without changing λ or y, so $y^f(\theta') = y^f(\theta)$. Moreover, (6.5) implies $t^f(\theta') = t^f(\theta)$, so $f(\theta') = f(\theta)$ (recall that $b(0) = 0$). Thus, the rectangular property holds, and secure implementation is achieved.

Example 2 illustrates the difference between implementation in *strictly* dominant strategies, and secure implementation. In Example 2, telling the truth is not a strictly dominant strategy, because an agent who gets no cake may still get no cake - and the same transfer of money - after a small change in his type. However, this does not prevent secure implementation, as long as the change in his type does not change anyone else's transfer. This is why φ_i must be constant. (If φ_i is not a constant then it can happen that $t^f(\theta') \neq t^f(\theta)$ even though $y^f(\theta') = y^f(\theta)$.)

If (6.3) does not hold, then the Groves-Clarke mechanism with constant φ_i will still be non-bossy. However, the rectangular property will be violated. Since one agent may consume all of the cake when (6.3) is violated, $u_i(f(\theta)) = u_i(f(\theta_i', \theta_{-i}))$ implies either $\theta_i' = \theta_i$ or $y_i^f(\theta_i', \theta_{-i}) = y_i^f(\theta) = 0$ or $y_i^f(\theta_i', \theta_{-i}) = y_i^f(\theta) = 1$. But this no longer ensures that the first order condition (6.4) holds when θ is replaced by θ′. Therefore, $f(\theta') \neq f(\theta)$ is possible. Intuitively, there can be bad Nash equilibria where one agent exaggerates his valuation of cake and receives all of it, while all the other agents report very low valuations and receive no cake. Notice that this example shows that, in general, non-bossiness and strategy-proofness together do not imply the rectangular property.

Example 3: Serial cost sharing.

The social decision is $y = (y_1, y_2, ..., y_n)$, where y_i is agent *i*'s consumption of divisible "cake". But unlike Example 2, now cake can be produced (using money as input). The cost function is $c(y) = C(\sum_i y_i)$, where *C* is strictly increasing, differentiable and convex. Each valuation function v_i is strictly increasing and strictly concave in y_i (but doesn't depend on y_j for $j \neq i$). Moulin and Shenker (1992) define *serial cost sharing* and show that this SCF is strategy-proof and can be Nash implemented by an indirect mechanism. In general, the two properties of Nash implementability and strategy-proofness together do not imply the rectangular property (which requires double implementation by the same mechanism). However, serial cost sharing does satisfy the rectangular property. Suppose $u_i(f(\theta^*)) = u_i(f(\theta_i, \theta_{-i}^*))$ for all $i \in I$. The definition of serial cost sharing implies $f(\theta^*) = f(\theta_i, \theta_{-i}^*)$ for all $i \in I$. This implies that if $y_i^f(\theta^*) > 0$ then $\partial v_i(y^f(\theta^*), \theta_i) / \partial y_i = \partial v_i(y^f(\theta^*), \theta_i^*) / \partial y_i$. If $y_i^f(\theta^*) = 0$, then $\partial v_i(y^f(\theta^*), \theta_i) / \partial y_i \leq C'(\sum_j y_j^f(\theta^*))$ and $\partial v_i(y^f(\theta^*), \theta_i^*) / \partial y_i \leq C'(\sum_j y_j^f(\theta^*))$. In either case, $f(\theta^*) = f(\theta)$, so serial cost-sharing is securely implementable. Notice that in this example, there is no need for any assumptions that rule out corner solutions.

7. Discrete Social Decisions

The previous section showed that surplus-maximizing social choice functions can be securely implemented in many quasi-linear environments with divisible public or private goods. In this section, we show that secure implementation is more difficult if the set of social decisions is discrete. Consider a quasi-linear environment as in Section 6, but now Y is a finite set. For convenience, $Y = \{0,1\}$, and $c(0) = c(1) = 0$. (The arguments can be adapted to any discrete Y.) We normalize so $v_i(0,\theta_i) = 0$ for all θ_i. Thus, agent i's preferences are characterized by $v_i(1,\theta_i)$, the value to him of social decision $y = 1$. Without loss of generality we may suppose $v_i(1,\theta_i) = \theta_i$ for all θ_i. We assume θ_i can be any real number.

Notice that if by chance $\sum_{i \in I} \theta_i = 0$, then both $y = 0$ and $y = 1$ are surplus maximizing. In this situation, it may be unreasonable to assume that the social choice rule is single-valued. Thus, we will allow f to be a multi-valued *social choice correspondence* (SCC). The definition of secure implementation when f is an SCC is the same as Definition 1. (Thus, we require "full" implementation in dominant strategy equilibria and Nash equilibria). Notice that for implementation in *strictly* dominant strategies, the issue of multi-valuedness would be moot because a strictly dominant strategy must be unique. However, in this paper we consider domination in the weak sense, and a given type of player may have several (weakly) dominant strategies. Moreover, even if each player has a unique dominant strategy, there may be multiple Nash equilibria (some of which are in dominated strategies). Secure implementation does not require a unique Nash equilibrium, but it does require that all Nash equilibrium outcomes are socially optimal (see Example 4 below).

We again use the notation $f(\theta)=(y^f(\theta),t_1^f(\theta),t_2^f(\theta)...,t_n^f(\theta))$, but now $y^f(\theta)$ and $t_i^f(\theta)$ are the *sets* of optimal decisions and transfers, respectively. The SCC *f* is *surplus maximizing* if $\sum_{i\in I}\theta_i<0$ implies $y^f(\theta)=\{0\}$, and $\sum_{i\in I}\theta_i>0$ implies $y^f(\theta)=\{1\}$. No restriction is imposed if $\sum_{i\in I}\theta_i=0$. For a mechanism $g{:}S\to A$, let $g(s)=(y^g(s),t^g(s))$ denote the outcome, where $y^g(s)$ is the chosen public project and $t^g(s)$ the profile of transfers.

To see that some interesting social choice correspondences can be securely implemented, even with a discrete set of public decisions, consider the following "veto rule".

Example 4: A veto rule.

Consider the following SCC. There are no transfers. The public decision $y=0$ is always socially optimal. The public decision $y=1$ is socially optimal if and only if $\theta_i\geq 0$ for all *i*. Intuitively, $y=0$ is a "status quo" outcome which is always socially acceptable, but the social project $y=1$ is acceptable to society if and only if *all* agents prefer it to the status quo. (With this interpretation, the SCC is the "individually rational" correspondence.) This SCC is securely implemented by the following mechanism. Each agent says 0 or 1. If all say 1, $y=1$ is implemented. If at least one agent says 0, then $y=0$ is implemented. Notice that the dominant strategy is to say 0 if $\theta_i<0$ and 1 if $\theta_i>0$. Both strategies are dominant if $\theta_i=0$. There are no "bad" Nash equilibria, because each agent can "veto" the outcome $y=1$. The "veto rule" is non-dictatorial: for each agent *i*, there is a profile θ such that agent *i* strictly prefers $y=1$, but the unique socially optimal decision is $y=0$. However, it does not maximize the surplus, because $y=0$ is socially acceptable even if $\theta_i>0$ for all *i*. We now show that in fact surplus maximization cannot be achieved in this environment. Notice that this negative result holds, even though we do not require budget balance (i.e., $\sum_{i\in I}t_i\neq 0$ is allowed).

Theorem 5. *Consider the quasi-linear environment with Y* = {0,1}*. There is no SCC which is both securely implementable and surplus-maximizing.*

Proof: Suppose *f* is surplus-maximizing. In order to obtain a contradiction, suppose it is securely implemented by a mechanism *g*.

Fix a type profile θ and choose $s_j \in DS_j^g(\theta_j)$ for each *j*. Surplus maximization implies that for any *i*, $y^g(s)=0$ if θ_i satisfies

(7.1) $\theta_i < \sum_{j \neq i} \theta_j$.

If θ_i satisfies

(7.2) $\theta_i > \sum_{j \neq i} \theta_j$

then $y^g(s)=1$. Moreover, if (7.1) holds, then any $s_i \in DS_i^g(\theta_i)$ must give agent *i* the same transfer, say $t_i^g(s)=t_i^0(s_{-i})$. (Otherwise, the strategy that gives the lowest transfer and the same public decision $y^g(s)=0$ could not be a dominant strategy). Similarly, if (7.2) holds, then any $s_i \in DS_i^g(\theta_i)$ must give agent *i* the same transfer, say $t_i^g(s)=t_i^1(s_{-i})$.

Suppose θ is such that $\sum_{i \in I} \theta_i > 0$. Define a new profile θ′ as follows. For $i \in \{1,2\}$, define $\theta_i' = -\sum_{j \neq i} \theta_j - \varepsilon < \theta_i$, where $\varepsilon > 0$. Let $\theta_i' = \theta_i$ for all $i > 2$. For each *i*, choose $s_i' \in DS_i^g(\theta_i')$. Clearly, $\sum_{i \in I} \theta_i' < 0$. Moreover, for $i \in \{1,2\}$, $\theta_i + \sum_{j \neq i} \theta_j' < 0$. For all *i*, we have chosen $s_i \in DS_i^g(\theta_i)$ and $s_i' \in DS_i^g(\theta_i')$. By surplus maximization, $y^g(s')=0$ and $y^g(s_i, s_{-i}')=0$ for $i \in \{1,2\}$. We now claim that, for $i \in \{1,2\}$, if agent *i*'s true type is θ_i then $s_i' \in DS_i^g(\theta_i')$ is a best response against s_{-i}'. Indeed, choosing s_i' would result in payoff $t_i^0(s_{-i}')$, because the social decision would be $y^g(s')=0$. But this is also what is obtained by choosing $s_i \in DS_i^g(\theta_i)$, because $y^g(s_i, s_{-i}')=0$. Therefore, s_i' is indeed a best response against s_{-i}' for $i \in \{1,2\}$ when his true type is θ_i. For all $i > 2$, $\theta_i' = \theta_i$ and $s_i' \in DS_i^g(\theta_i') = DS_i^g(\theta_i)$. Therefore, $s' \in N^g(\theta)$. But $y^g(s')=0$ even though $\sum_{i \in I} \theta_i > 0$, which contradicts the definition of surplus maximization. Q.E.D.

Notice that the proof of Theorem 5 in effect replicates the proof that the rectangular property is necessary for secure implementation, and then shows that the rectangular property is violated.

To further illustrate the impossibility of combining secure implementation with surplus maximization in the discrete environment, we consider a well-known example.

Example 5: Auctioning an indivisible object.

Suppose the social decision is to allocate a private *indivisible* object among two agents. Agent *i*'s true value of the object is $\theta_i \geq 0$ if she receives it, and 0 otherwise ($i = 1,2$). Consider the second price auction (Vickrey (1961)). Suppose $\theta_1 > \theta_2 > 0$. In order to maximize the surplus, agent 1 should win the object. Figure 1 shows that the set of Nash equilibria is quite large. The lower-right part of the set of Nash equilibria is the "good set" in the sense that agent 1 receives the object. However, the upper-left part of the set of Nash equilibria is "bad" in the sense that agent 2 receives the object, so the social surplus is not maximized.

Link to Figure 2
----

8. Single-Peaked Voting

Section 6 showed the possibility of secure implementation when the social decision is concerned with continuous variables, such as divisible public or private goods. However, the mechanisms relied on the existence of "money" for side-payments. We now show that if there are no side-payments, the results are negative, even if the social decision is a continuous variable.

Consider a *single-peaked voting environment*. The set of alternatives is $A = [0,1]$, and the set of possible preference relations consists of all those that are continuous and single-peaked on *A*. Let $p(u_i)$ denote the "peak" of u_i, i.e., the top ranked alternative in *A*, which is assumed to be unique. Single-peakedness implies that u_i is strictly increasing before $p(u_i)$ and strictly decreasing afterwards. Let $Range(f)$ denote the range of *f*. By Lemma 1 in Barberà and Jackson (1994), $Range(f)$ is closed. Let $a = \min\{x : x \in Range(f)\}$ and $b = \max\{x : x \in Range(f)\}$ denote the smallest and largest elements in $Range(f)$, respectively. Notice that *f* is constant if and only if *a* = *b*.

In the single peaked voting environment one can find dominant strategy implementable social choice functions with good properties, the leading example being the median voter rule (see Barbera and Jackson (1994)). This SCF is both non-dictatorial and Pareto efficient. Unfortunately, if a Pareto efficient social choice rule can be securely implemented, then it must be dictatorial. This is true even if we allow the social choice rule to be multi-valued. Before proving these negative results for secure implementation, we will prove two lemmas.

Lemma 6. *Let f be a securely implementable non-constant SCF in the single peaked voting environment. There is an agent i and an alternative* $y \in Range(f)$, $y > a$, *such that* $f(u) = y$ *whenever* $p(u_i) \geq y \geq p(u_j)$ *for all* $j \neq i$.

Proof: Let u' be any profile such that $p(u'_i) = a$ for all *i*, and let u'' be any profile such that $p(u''_i) = b$ for all *i*. Strategy-proofness implies $f(u') = a$ and $f(u'') = b$, and $b \neq a$ since *f* is not constant. If $f(u''_i, u'_{-i}) = a$ for all *i*, then the rectangular property implies that $f(u'') = a$, but this contradicts $f(u'') = b$. Thus, there is an agent, say agent *i* = 1, such that $f(u''_1, u'_{-1}) > a$. Define $y \equiv f(u''_1, u'_{-1})$.

Now let u_1 be any utility function such that $p(u_1) \geq y$. Consider $f(u_1, u'_{-1})$. If $f(u_1, u'_{-1}) > y$, then $u''_1(f(u_1, u'_{-1})) > u''_1(f(u''_1, u'_{-1}))$, and if $f(u_1, u'_{-1}) < y$, then $u_1(f(u''_1, u'_{-1})) > u_1(f(u_1, u'_{-1}))$. Since in either case we have a contradiction of strategy-proofness, we conclude that $f(u_1, u'_{-1}) = y$.

Now, for each $j \geq 2$, let u_j be any utility function such that $p(u_j) \leq y$. Consider $f(u_1, u_j, u'_{-1,j})$. If $f(u_1, u_j, u'_{-1,j}) > y$, then $u_j(f(u_1, u'_j, u'_{-1,j})) >$ $u_j(f(u_1, u_j, u'_{-1,j}))$, and if $f(u_1, u_j, u'_{-1,j}) < y$, then $u'_j(f(u_1, u_j, u'_{-1,j})) > u'_j(f(u_1, u'_j, u'_{-1,j}))$. Since in either case we have a contradiction of strategy-proofness, we conclude that $f(u_1, u_j, u'_{-1,j}) = y$ for all $j \geq 2$.

The rectangular property implies that $f(u) = y$. Thus, $f(u) = y$ whenever $p(u_1) \geq y \geq p(u_j)$ for all $j \geq 2$. Q.E.D.

Lemma 7. *Let f be a securely implementable non-constant SCF in the single peaked voting environment. There is an agent i such that* $f(u) = a$ *whenever* $p(u_i) = a$*, and* $f(u) = b$ *whenever* $p(u_i) = b$.

Proof: Without loss of generality, suppose agent $i = 1$ is the agent identified in Lemma 6, and y the alternative identified in the same lemma. Let u' be any profile such that $top(u'_i) = a$ for all i, and let $\tilde{u}$ be any profile such that $p(\tilde{u}_i) = y$ for all i. Then $f(u') = a$ by strategy-proofness, and Lemma 6 implies $f(\tilde{u}) = y$. If $f(u'_i, \tilde{u}_{-i}) = y$ for all i, then the rectangular property implies that $f(u') = y$, but this contradicts $f(u') = a$. Thus, there is an agent i such that $f(u'_i, \tilde{u}_{-i}) \neq y$. Lemma 6 implies that in fact $i = 1$. Strategy-proofness implies $f(u'_1, \tilde{u}_{-1}) < y$. Let $z \equiv f(u'_1, \tilde{u}_{-1}) < y$. We will show that $z = a$.

It is impossible that $z < a$ because $a = \min\{x : x \in Range(f)\}$. Suppose $z > a$. Now let $\hat{u}$ be a profile such that $p(\hat{u}_i) = z$ for all i. Strategy-proofness implies $f(\hat{u}) = z$. Since $z = f(u'_1, \tilde{u}_{-1})$, strategy-proofness implies $f(u'_1, \hat{u}_i, \tilde{u}_{-1,i}) = z$ for all $i > 1$. The rectangular property then implies $f(u'_1, \hat{u}_{-1}) = z$.

Now consider $f(u'_i, \hat{u}_{-i})$ for $i > 1$. Strategy-proofness requires $f(u'_i, \hat{u}_{-i}) \leq z$. Notice that this inequality holds regardless of $\hat{u}_1$, as long as $p(\hat{u}_1) = z$. Moreover, $f(u'_i, \hat{u}_{-i})$ is in fact the same alternative for any $\hat{u}_1$ such that $top(\hat{u}_1) = z$. (Otherwise, there would exist $\hat{u}_1$ and $\overline{u}_1$ such that $p(\hat{u}_1) = p(\overline{u}_1) = z$, and $f(\hat{u}_1, u'_i, \hat{u}_{-1,i}) <$ $f(\overline{u}_1, u'_i, \hat{u}_{-1,i}) \leq z$. But then $\hat{u}_1(f(\hat{u}_1, u'_i, \hat{u}_{-1,i})) < \hat{u}_1(f(\overline{u}_1, u'_i, \hat{u}_{-1,i}))$, contradicting strategy-proofness.) Suppose $w \equiv f(u'_i, \hat{u}_{-i}) < z < y$. But now consider $\hat{u}_1$ such that

$p(\hat{u}_1) = z$ and $\hat{u}_1(y) > \hat{u}_1(w)$. Lemma 6 implies that if $p(\tilde{u}_1) = y$, then $f(\tilde{u}_1, u_i', \hat{u}_{-1,i}) = y$. But since $\hat{u}_1(y) > \hat{u}_1(w)$ and $w = f(u_i', \hat{u}_{-i})$, strategy-proofness is violated. This contradiction implies $f(u_i', \hat{u}_{-i}) = z$ for all $i > 1$. Since we already have established $f(u_1', \hat{u}_{-1}) = z$, we can apply the rectangular property and conclude that $f(u') = z$. However, $f(u') = a$, a contradiction of our hypothesis that $z > a$. So, we must have $z = a$.

The previous paragraph established that $f(u_1', \tilde{u}_{-1}) = a$ whenever $p(u_1') = a$ and $p(\tilde{u}_i) = y$ for all $i > 1$. Now for all $i > 1$, let $\tilde{u}_i$ be such that $p(\tilde{u}_i) = y$, and $\tilde{u}_i(x) > \tilde{u}_i(a)$ for all $x \in Range(f)$ such that $x > a$. Consider any agent $i > 1$ and any arbitrary u_i. If $f(u_1', u_i, \tilde{u}_{-1,i}) \neq a$, then $\tilde{u}_i(f(u_1', u_i, \tilde{u}_{-1,i})) > \tilde{u}_i(f(u_1', \tilde{u}_i, \tilde{u}_{-1,i}))$, which contradicts strategy-proofness. Hence, $f(u_1', u_i, \tilde{u}_{-1,i}) = a$ for all $i > 1$. The rectangular property implies $f(u_1', u_{-1}) = a$. We conclude that $f(u_1', u_{-1}) = a$ whenever $p(u_1') = a$.

Exactly the same line of reasoning establishes the existence of an agent *i* such that $f(u_i', u_{-i}) = b$ whenever $p(u_i') = b$. Obviously, this must be *i* = 1, or else we contradict the already established fact that $f(u_1', u_{-1}) = a$ whenever $p(u_1') = a$. Q.E.D.

Now we are ready to prove our first negative result for single peaked voting. It covers the case of single-valued social choice rules.

Theorem 7. *Let f be a securely implementable SCF in the single peaked voting environment. There is a dictator on $Range(f)$, i.e., an agent i such that for all u and all $x \in Range(f)$, $u_i(f(u)) \geq u_i(x)$.*

Proof: Since the result is trivial if *f* is constant, suppose *f* is securely implementable but not constant. Lemma 7 identifies an agent *i* such that $f(u) = a$ whenever $p(u_i) = a$, and $f(u) = b$ whenever $p(u_i) = b$. Without loss of generality suppose this is true for *i* =

1. Fix any $x \in Range(f)$. Let u' be such that $p(u_i') = x$ for all i, and let u be an arbitrary profile. The theorem is proved by showing that $f(u_1', u_{-1}) = x$ must necessarily hold.

Strategy-proofness implies $f(u') = x$. Fix any $i > 1$. We will show that $f(u_i, u_{-i}') = x$. If $p(u_i) = x$, then $f(u_i, u_{-i}') = x$ by strategy-proofness. Suppose instead that $p(u_i) > x$. Then strategy-proofness implies $f(u_i, u_{-i}') \geq x$. Notice that this inequality holds regardless of u_1', as long as $p(u_1') = x$. Moreover, $f(u_i, u_{-i}')$ is in fact the same alternative for any u_1' such that $p(u_1') = x$. (Otherwise, there would exist u_1' and u_1'' such that $p(u_1') = p(u_1'') = x$, and $f(u_1', u_i, u_{-1,i}') > f(u_1'', u_i, u_{-1,i}') \geq x$. But then $u_1'(f(u_1', u_i, u_{-1,i}')) < u_1'(f(u_1'', u_i, u_{-1,i}'))$, contradicting strategy-proofness.)

Now suppose $z \equiv f(u_i, u_{-i}') > x$. But consider u_1' such that $p(u_1') = x$, and $u_1'(a) > u_1'(z)$. If $\tilde{u}_1$ is such that $p(\tilde{u}_1) = a$, then $f(\tilde{u}_1, u_i, u_{-1,i}') = a$ by Lemma 8. But then $u_1'(f(\tilde{u}_1, u_i, u_{-1,i}')) > u_1'(f(u_1', u_i, u_{-1,i}'))$, contradicting strategy-proofness. This contradiction shows that we must have $f(u_i, u_{-i}') = x$ whenever $p(u_i) > x$. A similar argument establishes that $f(u_i, u_{-i}') = x$ whenever $p(u_i) < x$. We conclude that, for all $i > 1$, $f(u_i, u_{-i}') = x$ for all u_i. The rectangular property implies $f(u_1', u_{-1}) = x$. Q.E.D.

As in the previous section, there exist non-dictatorial social choice correspondences that can be securely implemented. For example, a "veto rule", similar to Example 4, with some arbitrary alternative designated as status quo, can be securely implemented in the single-peaked voting model. However, this SCC is not Pareto efficient. More generally, in this environment an SCC is either single-valued, in which case it must be dictatorial by Theorem 7, or it is Pareto inefficient. This is our second negative result for single-peaked voting.

Theorem 8. *Let f be a securely implementable SCC in the single peaked voting environment. Then f is either single-valued or Pareto inefficient.*

Proof: Suppose f is a securely implementable SCC which is not single-valued. Then there is u such that $f(u)$ contains at least two distinct alternatives. If f is securely implemented by mechanism g, then there must be two strategy profiles $s, s' \in DS^g(u)$ such that $g(s) \neq g(s')$. Then, there must necessarily exist alternatives a and b, and an agent i, such that $g(s'_1, \ldots, s'_{-i}, s_i, s_{i+1}, \ldots, s_n) = a$ but $g(s'_1, \ldots, s'_{-i}, s'_i, s_{i+1}, \ldots, s_n) = b \neq a$. We may choose labeling so that $i = 1$, and $b > a$.

Thus, we have $s \in DS^g(u)$, $g(s) = a$, $(s'_1, s_{-1}) \in DS^g(u)$ and $g(s'_1, s_{-1}) = b > a$. Since $s_1, s'_1 \in DS_1^g(u_1)$, it must be the case that $a < p(u_1) < b$ and $u_1(a) = u_1(b)$.

Let $L = \{j : p(u_j) \leq a\}$ be the set of agents whose peaks, in the profile u, are (weakly) to the left of a. Suppose $2 \in L$, and suppose u_2^* is such that $a < p(u_2^*) < b$ and $u_2^*(a) > u_2^*(b)$. Let $s_2^* \in DS_2^g(u_2^*)$.

Claim: $g(s_1, s_2^*, s_{-1,2}) = a$.

To prove the claim, we consider the various possibilities.

Case 1: $a < g(s_1, s_2^*, s_{-1,2}) < b$. Since $s_1, s'_1 \in DS_1^g(u_1)$, we must have $u_1(g(s_1, s_2^*, s_{-1,2})) = u_1(g(s'_1, s_2^*, s_{-1,2}))$. Therefore, $a < g(s'_1, s_2^*, s_{-1,2}) < b$. But $g(s'_1, s_2, s_{-1,2}) = b$ and $2 \in L$, so $u_2(g(s'_1, s_2^*, s_{-1,2})) > u_2(g(s'_1, s_2, s_{-1,2}))$. However, this contradicts $s_2 \in DS_2^g(u_2)$. Therefore, case 1 is impossible.

Case 2: $g(s_1, s_2^*, s_{-1,2}) < a = g(s)$. This case is impossible because $p(u_2^*) > a$ and $s_2^* \in DS_2^g(u_2^*)$.

Case 3: $g(s_1, s_2^*, s_{-1,2}) \geq b$. But then, $u_2^*(g(s)) > u_2^*(b) \geq u_2^*(g(s_1, s_2^*, s_{-1,2}))$, which contradicts $s_2^* \in DS_2^g(u_2^*)$.

Since cases 1,2 and 3 are all impossible, the claim is true.

The claim establishes $g(s_1, s_2^*, s_{-1,2}) = a$. Since $s_1, s'_1 \in DS_1^g(u_1)$, it must be the case that $u_1(g(s'_1, s_2^*, s_{-1,2})) = u_1(a)$. This means that $g(s'_1, s_2^*, s_{-1,2})$ can be either a or b. Suppose $g(s'_1, s_2^*, s_{-1,2}) = a$. But, $g(s'_1, s_{-1}) = b$. Since $2 \in L$ we have $u_2(a) > u_2(b)$, which contradicts $s_2 \in DS_2^g(u_2)$. Therefore, we must have $g(s'_1, s_2^*, s_{-1,2}) = b$.

To summarize, we have shown that $(s_2^*, s_{-2}) \in DS^g(u_2^*, u_{-2})$, $g(s_2^*, s_{-2}) = a$, $(s_1', s_2^*, s_{-1,2}) \in DS^g(u_2^*, u_{-2})$ and $g(s_1', s_2^*, s_{-1,2}) = b > a$. This puts us back in our original position, except that the L set has one fewer member after u_2 is replaced by u_2^* (because $p(u_2^*) > a$). We can repeat the same argument for each $j \in L$: we let u_j^* be such that $a < p(u_2^*) < b$ and $u_j^*(a) > u_j^*(b)$, and we pick $s_j^* \in DS_j^g(u_j^*)$. After having exhausted all the members of L, we obtain $s_L^* = \{s_j^*\}_{j \in L}$, where $s_j^* \in DS_j^g(u_j^*)$ for each $j \in L$, and $g(s_{-L}, s_L^*) = a$. Since g securely implements f, $a \in f(u_{-L}, u_L^*)$. However, by definition of L, when the utility profile is (u_{-L}, u_L^*), all agents have peaks strictly to the right of a. Therefore, a is not Pareto efficient. Q.E.D.

9. 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 two agents and a binary public project that has a continuum of Nash equilibria, and a Groves-Clarke mechanism with two agents and single-peaked preferences that has a unique Nash equilibrium. We found that subjects played dominant strategies significantly more frequently in the secure Groves mechanism than in the non-secure pivotal mechanism. This makes us optimistic about the future of mechanism design. The negative experimental evidence mentioned in the introduction was based on mechanisms that are not secure (such as the second price auction). In these experiments, there may have been insufficient pressure on the players to adopt their dominant strategies, and deviations may not have been punished by big payoff losses (for a discussion, see Cason, Saijo, Sjöström and Yamato (2003)). Imposing stricter requirements than simply strategy-proofness may turn out to be the key to successful applications of mechanism design.

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. (1994) "A Characterization of Strategy Proof Social Choice Functions for Economies with Pure Public Goods", *Social Choice and Welfare,* 11, 241-252.

Bergemann, D. and Morris, S. (2003) "Robust Implementation: The Role of Large Type Spaces," mimeo.

Bergemann, D. and Morris, S. (2004) "Robust Mechanism Design," Cowles Foundation Discussion Paper No. 1421R, Yale University.

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.

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

Holmström, B. (1979). "Groves Scheme on Restricted Domain," *Econometrica*, 47, 1137-1144.

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.

Mizukami, H. and Wakayama, T. (2004) "Dominant Strategy Implementation: A Characterization," mimeo.

Moulin, H. and S. Shenker (1992) "Serial Cost Sharing," *Econometrica*, 60, 1009-1037.

Myerson, R. (1981) "Optimal Auction Design," *Mathematics of Operation Research*, 6, 58--73.

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

Satterthwaite, M. and Sonnenschein, H. (1981) "Strategy Proof Allocation Mechanisms at Differentiable Points," *Review of Economic Studies*, 48, 587-97.

Sprumont, Y. (1991) “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.

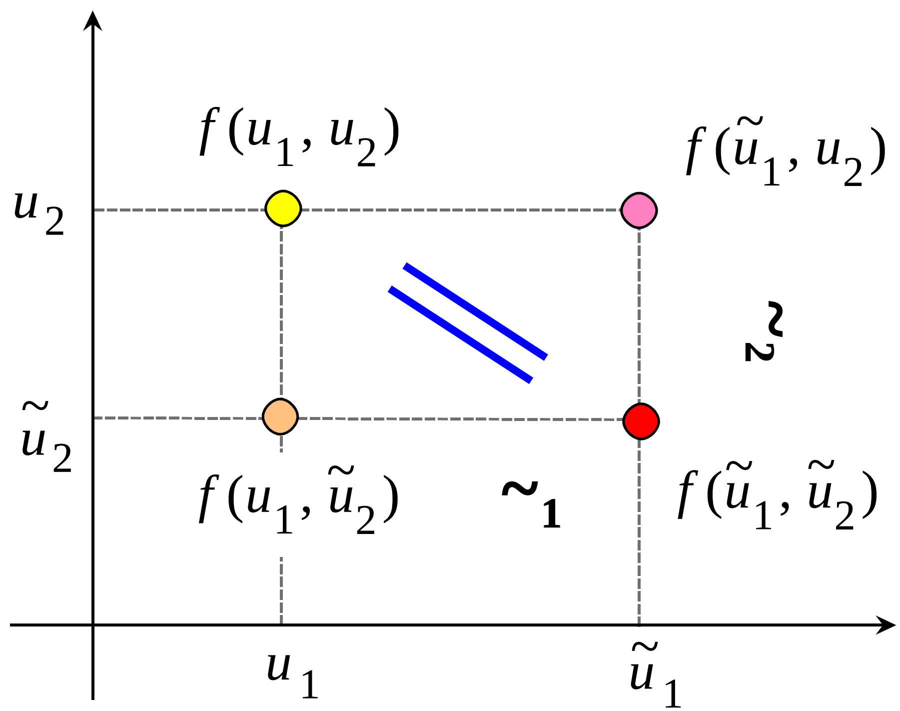

Figure 1: Rectangular Property

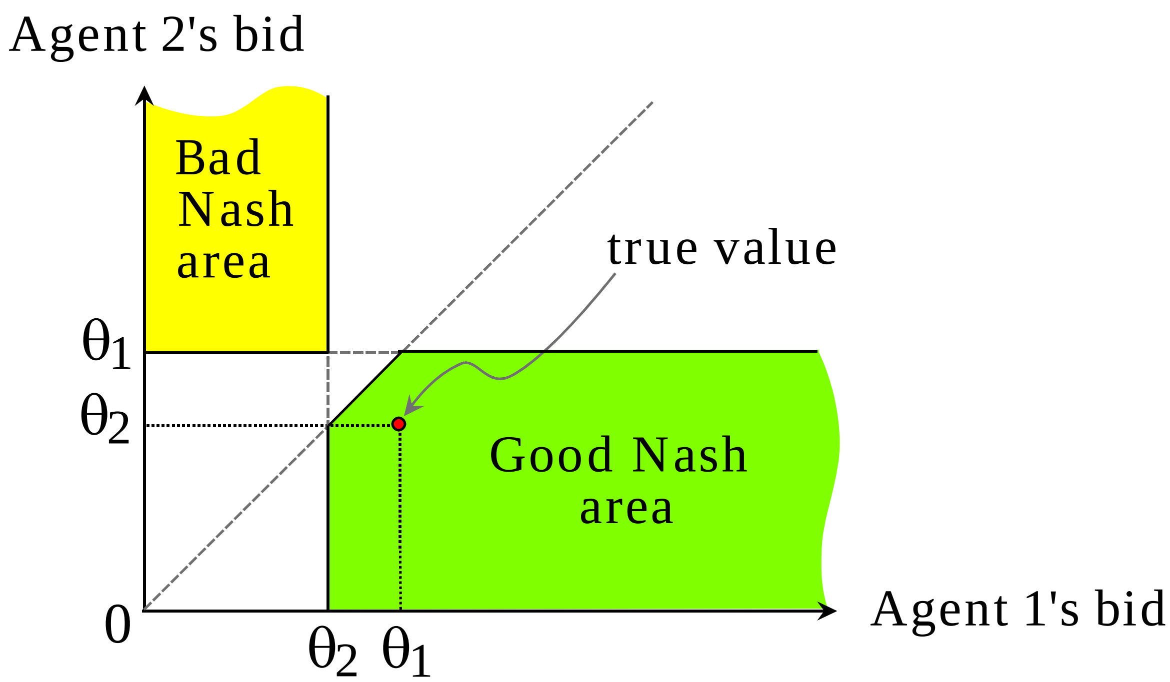

Figure 2: Equilibria of the Second Price Auction