The Korean Economic Review
Volume 29, Number 2, Winter 2013, 429-445.

A Sufficient Condition for Bayesian Implementation with Side Payments

Bong-Ju Kim*

It is well-known that with three agents and more, Bayesian monotonicity and the self-selection condition are both necessary and sufficient for Bayesian implementation in a general environment. Matsushima (1993) identified a condition, Condition 1, which is equivalent to Bayesian monotonicity under the strict self-selection (SSS) condition in an environment with side payments. This paper identifies a condition that is equivalent to Bayesian monotonicity under the self-selection condition in an environment with side payments when the prior probability is independently distributed. This condition identified here is weak in that its combination with the self-selection condition is weaker than the SSS condition.

JEL Classification: C72, D71, D82
Keywords: Bayesian Implementation, Side Payments, Strict Self-selection, Bayesian Monotonicity

I. Introduction

Bayesian implementation considers the problem of implementing a social choice function in an environment with incomplete information among the agents. Since the socially desirable outcome depends on agents' private information, it may be possible that agents do not have the incentive to correctly reveal their private information. This need to give agents the right incentives acts as a constraint both on the kind of decentralized procedures which can be used as well as on the class of socially desired outcome which can be implemented. Bayesian incentive compatibility of a social choice function is simply the requirement that each agent

Received: Oct. 17, 2012. Revised: Feb. 15, 2013. Accepted: March 4, 2013.
* National Assembly Research Service, 1 Uisadang-ro, Yeouido-dong Yeongdeungpo-gu, Seoul, 150-703, South Korea. Tel: 82-2-788-4590, Fax: 82-2-788-4599, E-mail: kbongju@gmail.com. I am grateful to the editor and referees for comments that led to significantly improvements of this paper. I am deeply indebted to Sungwhee Shin and Wanjin Kim for their guidance and encouragement. I am also grateful to Seung-Nyeon Kim and Soo Yeon Park for helpful comments.

has the incentive to truthfully reveal his or her information when all other agents report their information truthfully. More accurately, a social choice function is said to be incentive compatible if truth-telling is a Bayesian Nash equilibrium of the direct mechanism in which agents report their private information and the outcome is the social choice function to the socially desired outcome. The revelation principle states that if an allocation is incentive compatible, then it can be produced as the equilibrium outcome to the direct mechanism. By this principle, incentive compatibility of a social choice function is necessary for it to be implemented through a Bayesian Nash equilibrium of any mechanism. However, incentive compatibility is only half of the implementation problem. A mechanism applied to an incentive compatible social choice function may possess other equilibria (which do not correspond to the socially desired outcome). Full implementation refers to designing a mechanism that resolves this multiplicity problem by ensuring that all equilibria correspond the socially desired outcome in each information state, and requires some condition to incentive compatibility. For brevity, we drop the word "full" and simply refer to it here as "implementation."

The Bayesian approach to implementation with incomplete information was initiated by Postlewaite and Schmeidler (1986). They characterized implementable social choice set for exchange economies in which information is nonexclusive[1] and there are at least three agents. Under the nonexclusive information, a Bayesian monotonicity condition is necessary and sufficient for implementation in an economy, provided that there are at least three agents. However, nonexclusive information is restrictive. Their information restriction excludes many models of interest such as those employed to study auctions, public good provision, and optimal trading mechanisms with incomplete information. In these models and in most others, agents hold exclusively private information.

Palfrey and Srivastava (1989a) examined implementation for exchange economies in which agents may have exclusive information. They showed that an implementable collection of social choice functions must satisfy a Bayesian monotonicity condition and an incentive compatibility condition. They also showed that the Bayesian monotonicity and a slightly stronger incentive compatible condition are sufficient for implementation when there are at least three agents. In economic environments[2] with three or more agents, Jackson (1991) showed that a collection of social choice functions is implementable if and only if closure,

[1] Nonexclusive information means that each agent's information is redundant given the collective information of the other agents. Under this assumption, an agent's report of the state may be checked against the joint report of the other agents. Complete information is the extreme form of nonexclusive information.

[2] The definition of economic environments is that any given choice function and state, there are at least two agents who prefer to alter social choice function at that state.

incentive compatibility, and Bayesian monotonicity conditions are satisfied.[3] Extending the analysis to noneconomic environments, he showed that when there are at least three agents, closure, incentive compatibility,[4] and a combination of monotonicity and no-veto conditions[5] are sufficient for implementation. In general, however, it is not easy to check whether a social choice function satisfies the condition of Bayesian monotonicity. For restricted environments where agents' preferences are represented by quasi-linear utility functions and side payments are possible, Matsushima (1993) introduced an easily verifiable condition, so called Condition 1. This means that the set of all deceptions over which Bayesian monotonicity restrictions have to be verified is the set of all consistent deceptions. Condition 1 is easier to verify than Bayesian monotonicity because Bayesian monotonicity restrictions have to be verified over the set of all consistent deceptions.[6] He shows that Condition 1 is equivalent to Bayesian monotonicity under strict self-selection condition (SSS). Moreover, Matsushima (1993) showed that the following informational condition is sufficient for Condition 1: That is, there is no consistent deception other than truthful revelation which is called no-consistent deception condition. He showed that with three agents or more, any social choice function with SSS is implementable under the no-consistent deception condition. Assuming that agent types are independently distributed, Palfrey and Srivastava (1993) showed that if no-consistent deception is satisfied and there is a money good, then a social choice function is incentive compatible if and only if it is Bayesian implementable.

Following in the same vein as the two studies above, we identify an easily verifiable condition that is equivalent to Bayesian monotonicity under the self-selection condition in a more restricted environment with side payment and private values.[7] Assuming that the distribution of types is independent, we identify a condition in this environment equivalent to Bayesian monotonicity under the self-selection condition. The condition is weak in that its combination with the self-selection condition is weaker than the strict self-selection (SSS) condition in

[3] The closure condition requires that the social choice set be closed under concatenation of common knowledge events.

[4] The incentive compatibility is also called the self-selection. Henceforth we use self-selection instead of incentive compatibility.

[5] This combined condition is called Monotonicity-No-Veto (MNV). In fact MNV reduces to the separate condition monotonicity and No Veto Power when information is complete.

[6] The consistent deception generates the same distribution of reports by the informed agents as if they had all used their truthful strategies.

[7] The assumption of private values means that the utility of agent i depends on the signal he or she receives but not on the signals that other agents receive. In contrast, Matsushima (1993) considers a more general case, the so-called common values case, where the utility of agent i depends on the signals that all the agents receive. Thus, our private values environment is a special case of that of Matsushima (1993).

Matsushima (1993). This combination is termed the weakly strict self-selection (WSSS) condition. Note that WSSS condition is a sufficient condition, under quasi-linear utility function with side payments, for Bayesian implementation. Thus, Condition 1 in Matsushima (1993) is dispensed with in an environment with side payment and private values. This means that the satisfaction of a stronger self-selection condition, that is, WSSS is sufficient for Bayesian implementation.

II. The Model

Let $N = \{1, \ldots, n\}$ denote the finite set of agents and S_i be the finite set of types or signals of agent $i \in N$. A profile of signals defines a state $s = (s_1, \ldots, s_n)$ and $S = \times_{i \in N} S_i$ denotes the set of possible states.

Let $p_i(s)$ denote agent i's prior probability that agents receive the profile s of signals. Let s_{-i} denote the set of signals received by other agents than the agent i. Let $p_i(s_{-i} \mid s_i)$ denote the conditional probability that other agents receive the signals s_{-i} when the agent i receives the signal s_i. We assume that the conditional probability is independently distributed, that is, $p_i(s_{-i} \mid s_i) = p_i(s_{-i} \mid s_i')$ for all $i \in N$, $s_i, s_i' \in S_i$. Henceforth, for brevity, we let $p_i(s_{-i})$ denote the probability that does not depend on $s_i \in S_i$ for all $i \in N$. We assume that for every $i \in N$ and every $s \in S$, $p_i(s_{-i}) > 0$.[8] Let A denote the set of lotteries over some arbitrary finite set of alternatives. We assume unrestricted side payments with full transferability.

Agent i's state-dependent von Neumann-Morgenstern utility function is of the form $U_i(a, r_i, s_i) = u_i(a, s_i) + r_i$ where r_i is the transfer to agent i, and $u_i(a, s_i)$ is the (expected) utility which agent i receives from the lottery $a \in A$ in the absence of any transfer payment.

A public decision rule $x : S \to A$ maps from states to lotteries. A transfer rule $t = \{t_i\}_{i \in N} : S \to R^n$ maps from states to an n-tuple of real numbers. A social choice function is a pair (x, t) composed by a public decision rule and a transfer rule. Agent i's conditional expected utility from a social choice function (x, t) when he receives the signal s_i is

$$\sum_{s_{-i} \in S_{-i}} U_i(x(s), t(s), s_i) p_i(s_{-i}) .$$

Let $\phi_i : S_i \to S_i$ denote a deception of agent $i \in N$ and Φ_i the set of all deceptions of agent $i \in N$. Let $\phi = (\phi_i)_{i \in N}$ denote a profile of individual deceptions

[8] We assume that information is diffuse in the sense that no strict subset of agents can pool their information and rule out certain types of the other agents.

and $\Phi = \times_{i \in N} \Phi_i$ denote the set of all profiles of individual deceptions. We will call ϕ simply a deception.

For every $i \in N$, every $\phi \in \Phi$ and every $s \in S$, we define

$$\hat{p}_i(s_{-i} \mid s_i, \phi_{-i}) = \sum_{s'_{-i} \in \phi^{-1}_{-i}(s_{-i})} p_i(s'_{-i}) [= \sum_{s'_{-i} \in \phi^{-1}_{-i}(s_{-i})} p_i(s'_{-i} | s_i)].$$

$\hat{p}_i(s_{-i} \mid s_i, \phi_{-i})$ is the probability of observing report s_{-i} in a direct mechanism given agent i's signal s_i when the other agents use the deception ϕ_{-i}.

Now after introducing the concept of Consistent deception, we introduce that of No Consistent deception.

Consistent deception: A deception ϕ is consistent if and only if for every $i \in N$ and every $s \in S$, $p_i(s_{-i}) = [p_i(s_{-i} \mid \phi_i(s_i))] = \hat{p}_i(s_{-i} \mid s_i, \phi_{-i})$.

This means that for every $i \in N$ and every $s \in S$, the probability of observing report s_{-i} given agent i's signal $\phi_i(s_i)$ when agents are truthful, i.e., $p_i(s_{-i}) [= p_i(s_{-i} \mid \phi_i(s_i))]$, is equivalent to the probability of observing report s_{-i} given agent i's signal s_i when ϕ_{-i} is used, i.e., $\hat{p}_i(s_{-i} \mid s_i, \phi_{-i})$.

No-consistent deception (NCD): NCD means that there is no consistent deception other than truthful revelation, i.e., $\phi(s) = s$.

Now we introduce some conditions on the social choice function (x, t).

Bayesian monotonicity: A social choice function (x,t) satisfies Bayesian monotonicity if it satisfies the following condition:

$$\text{For every deception } \phi \text{, if } (x,t)(\phi(s)) \neq (x,t)(s) \text{ for some } s \in S, \tag{1}$$

then there exists $i \in N$, $s_i \in S_i$, and another allocation rule (x', t') such that

$$\sum_{s_{-i} \in S_{-i}} [U_i(x'(\phi(s)), t'_i(\phi(s)), s_i)] p_i(s_{-i}) > \sum_{s_{-i} \in S_{-i}} [U_i(x(\phi(s)), t_i(\phi(s)), s_i)] p_i(s_{-i}), \tag{2}$$

and for every $s''_i \in S_i$,

$$\begin{aligned} &\sum_{s''_{-i} \in S_{-i}} [U_i(x(s''), t_i(s''), s''_i)] p_i(s''_{-i}) \\ &\geq \sum_{s''_{-i} \in S_{-i}} [U_i(x'(s''_{-i}, \phi_i(s_i)), t'_i(s''_{-i}, \phi_i(s_i)), s''_i)] p_i(s''_{-i}). \end{aligned} \tag{3}$$

As mentioned in Jackson (1991), the Bayesian monotonicity condition assures the selective elimination of undesirable equilibria. Before suggesting intuitive interpretation for this condition, we introduce some definitions. A mechanism is a pair (M,g) where M_i is agent i's message space, $M=\prod_{i=1}^{n} M_i$ is the product of the individual message spaces and $g:M\rightarrow A$ is an outcome function. A strategy for agent i is a mapping $\sigma_i : S_i \rightarrow M_i$. Let σ denote the vector of strategies $\sigma=[\sigma_1,\ldots,\sigma_n]$. the Bayesian monotonicity condition means that if $(x,t)(\phi(s))\neq(x,t)(s)$ for some $s\in S$, then $\sigma\circ\phi$ must not be an equilibrium. The Bayesian monotonicity assures that in this case $\sigma\circ\phi$ can be ruled out as an equilibrium. The existence (x',t') with the stated properties allows agent i to signal that ϕ is being played. Inequality (2) assures that the agent is rewarded according to (x',t'), which makes him or her better off. The other part of the Bayesian monotonicity condition, i.e., inequality (3) assures that agent i cannot gain by falsely accusing the other agents of deceiving. Notice that assuming independent agents' types, $p_i(s_{-i}\,|\,s_i)=p_i(s_{-i})$ and $p_i(s''_{-i}\,|\,s''_i)=p_i(s''_{-i})$.

Now after introducing the concept of self-selection and strict self-selection, we introduce that of Consistent S.

Self-Selection (SS): For all $i\in N$, all s_i, $s'_i\in S$,

$$\sum_{s_{-i}\in S_{-i}} U_i(x(s),t_i(s),s_i)p_i(s_{-i})\geq \sum_{s_{-i}\in S_{-i}} U_i(x(s_{-i},s'_i),t_i(s_{-i},s'_i),s_i)p_i(s_{-i})\,. \tag{4}$$

Strict Self-Selection (SSS): For all $i\in N$, all $s_i\in S$, and all $s'_i\in S\setminus\{s\}$,

$$\sum_{s_{-i}\in S_{-i}} U_i(x(s),t_i(s),s_i)p_i(s_{-i})> \sum_{s_{-i}\in S_{-i}} U_i(x(s_{-i},s'_i),t_i(s_{-i},s'_i),s_i)p_i(s_{-i})\,. \tag{5}$$

SS is the property that, in the direct mechanism, each agent prefers reporting his or her type truthfully to misrepresenting it, as long as all other agents also truthfully report their types. In this case, however, SSS requires that each agent strictly prefer reporting his or her type truthfully to misrepresenting it.

Condition S: If a social choice function (x,t) satisfies inequality (1) for a consistent deception ϕ, then

$$\sum_{s_{-i}\in S_{-i}} U_i(x(s),t_i(s),s_i)p_i(s_{-i})> \sum_{s_{-i}\in S_{-i}} U_i(x(s/\phi_i(s_i)),t_i(s/\phi_i(s_i)),s_i)p_i(s_{-i})$$

for some $i\in N$ and $s_i\in S_i$. (6)

Condition S means that SSS restrictions have to be verified over the set of consistent deception with inequality (1). In addition, for the consistent deception this condition requires that there exist an agent i who uses the deception $\phi_i(s_i) \neq s_i$ and strictly prefers reporting his or her type truthfully to misrepresenting it.

Note that the combination of SS and Condition S is weaker than SSS condition. The combination of SS and Condition S is called the weakly strict self-selection (WSSS) condition. This means that WSSS condition is stronger than the self-selection condition, but it is weaker than SSS condition.

To compare Condition S with Condition 1 in Matsushima (1993), we introduce the latter adapted to our current notation as follows:

Condition 1: For every consistent deception ϕ, if (x,t) satisfies inequality (1), then

$$\sum_{s_{-i}\in S_{-i}} [U_i(x'(\phi(s)), t_i'(\phi(s)), s_i) - U_i(x(\phi(s)), t_i(\phi(s)), s_i)] p_i(s_{-i})$$
$$> \sum_{s'_{-i}\in S_{-i}} [U_i(x'(s'_{-i}, \phi_i(s_i)), t_i'(s'_{-i}, \phi_i(s_i)), \phi_i(s_i))$$
$$- U_i(x(s'_{-i}, \phi_i(s_i)), t_i(s'_{-i}, \phi_i(s_i)), \phi_i(s_i))] p_i(s'_{-i}), \text{ for some } i \in N, \text{ and } s_i \in S_i,$$
and another allocation rule (x', t'). (7)

Inequality (7) implies that the difference between agent i's conditional expected direct returns from x and from x' under the signal s_i when ϕ is used is larger than that under the signal $\phi_i(s_i)$ when all agents are truthful. Notice that assuming independent agents' types, $p_i(s_{-i}|s_i) = p_i(s_{-i})$ and $p_i(s'_{-i}|\phi_i(s_i)) = p_i(s'_{-i})$.

III. Main Results

In the following Proposition, we directly prove the logical relation between the WSSS condition and Bayesian monotonicity.

Proposition 1. *Suppose that the agents' types are independently distributed. Moreover, suppose that a social choice function satisfies SS. The social choice function satisfies Condition S if and only if it satisfies Bayesian monotonicity.*

Proof. First, we show that Condition S implies Bayesian monotonicity. Consider a deception that satisfies (1). We consider two cases separately for the possible forms of deceptions; i.e., consistent deceptions and no consistent deceptions.

Case 1 (consistent deception):
Consider the social choice function (x',t') such that, irrespective of the agent i's report, given a state of $s_i \in S_i$, the social choice function $(x',t')(s') = (x,t)(s'_{-i}, s_i)$ for all $s' \in S$.

(i) We show that (3) holds.

We know that the right hand side (RHS) of inequality (3) is agent i's expected utility corresponding to the social choice function $(x',t')(s'') = (x,t)(s''_{-i}, s_i)$ for all $s'' \in S$. Moreover, we know that the left hand side (LHS) of inequality (3) represents agent i's expected utility when all agents report their type truthfully. By the SS condition, the LHS of inequality (3) is not less than the RHS of (3).

(ii) We show that the inequality (2) holds for some i and $s_i \in S_i$.

The inequality (2) is equivalent to the following:

$$\sum_{s_{-i} \in S_{-i}} [U_i(x'(\phi(s)), t'_i(\phi(s)), s_i) - U_i(x(\phi(s)), t_i(\phi(s)), s_i)] p_i(s_{-i}) > 0 . \tag{8}$$

It can be shown that (8) holds as follows:

$$\begin{aligned}
&\sum_{s_{-i} \in S_{-i}} [U_i(x'(\phi(s)), t'_i(\phi(s)), s_i) - U_i(x(\phi(s)), t_i(\phi(s)), s_i)] p_i(s_{-i}) \\
&= \sum_{s'_{-i} \in S_{-i}} [U_i(x'(s'_{-i}, \phi_i(s_i)), t'_i(s'_{-i}, \phi_i(s_i)), s_i) \\
&\quad - U_i(x(s'_{-i}, \phi_i(s_i)), t_i(s'_{-i}, \phi_i(s_i)), s_i)] \hat{p}_i(s'_{-i} \mid s_i, \phi_{-i}) \\
&= \sum_{s'_{-i} \in S_{-i}} [U_i(x'(s'_{-i}, \phi_i(s_i)), t'_i(s'_{-i}, \phi_i(s_i)), s_i) \\
&\quad - U_i(x(s'_{-i}, \phi_i(s_i)), t_i(s'_{-i}, \phi_i(s_i)), s_i)] p_i(s'_{-i} \mid \phi_i(s_i)) \\
&= \sum_{s'_{-i} \in S_{-i}} [U_i(x(s'_{-i}, s_i), t_i(s'_{-i}, s_i), s_i) - U_i(x(s'_{-i}, \phi_i(s_i)), t_i(s'_{-i}, \phi_i(s_i)), s_i)] p_i(s'_{-i}) \\
&> 0 .
\end{aligned} \tag{9}$$

The first equality follows from the substitution of $\phi_{-i}(s_{-i})$ with s'_{-i} and the definition of $\hat{p}_i(s_{-i} \mid s_i, \phi_{-i})$. The second equality follows from the consistency of ϕ. The third equality follows from the definition of (x',t') and the independence of prior probability. The last inequality follows from the Condition S.

Case 2 (inconsistent deception):
Since ϕ is not consistent, then there exist $i \in N$ and $s \in S$ such that

$$p_i(s_{-i})[=p_i(s_{-i}\,|\,\phi_i(s_i))]\neq \hat{p}_i(s_{-i}\,|\,s_i,\phi_{-i}).$$

In the case of independent types, this ensures that there exists a function $\mu : S_{-i} \to R$ such that:

$$\sum_{s'_{-i}\in S_{-i}} \mu(s'_{-i})p_i(s'_{-i}) \leq 0, \tag{10}$$

$$\sum_{s'_{-i}\in S_{-i}} \mu(s'_{-i})\hat{p}_i(s'_{-i}\,|\,s_i,\phi_{-i}) > 0. \tag{11}$$

Define a transfer rule t' such that for every $s\in S$, $t'_i(s)=t_i(s)+\mu(s_{-i})$, $t'_{i+1}(s)=t_{i+1}(s)-\mu(s_{-i})$, and $t'_j(s)=t_j(s)$ for all $j\in N/\{i,i+1\}$. It is clear that (x,t') satisfies inequalities (2) and (3) of Bayesian monotonicity. Since ϕ is not consistent, $\mu : S_{-i} \to R$ can be designed so that whenever a deception is used, at least one agent obtains gains for proposing a non-zero transfer.

Next, we show that Bayesian monotonicity implies Condition S. Suppose that (x,t) satisfies Bayesian monotonicity. In addition, suppose that a consistent deception ϕ satisfies inequality (1). It is clear from Bayesian monotonicity that there exists $i\in N$, $s_i\in S_i$, and another allocation rule (x',t') that satisfies inequalities (2) and (3). By the independence of prior probability and the consistency of ϕ, (2) is simplified as follows:

$$\sum_{s''_{-i}\in S_{-i}} U_i(x'(s''_{-i},\phi_i(s_i)),t'_i(s''_{-i},\phi_i(s_i)),s_i)p_i(s''_{-i})$$
$$> \sum_{s''_{-i}\in S_{-i}} U_i(x(s''_{-i},\phi_i(s_i)),t_i(s''_{-i},\phi_i(s_i)),s_i)p_i(s''_{-i}). \tag{12}$$

From (3), when $s''_i=s_i$, we have the following inequality.

$$\sum_{s''_{-i}\in S_{-i}} U_i(x(s''_{-i},s_i),t_i(s''_{-i},s_i),s_i)p_i(s''_{-i})$$
$$\geq \sum_{s''_{-i}\in S_{-i}} U_i(x'(s''_{-i},\phi_i(s_i)),t'_i(s''_{-i},\phi_i(s_i)),s_i)p_i(s''_{-i}). \tag{13}$$

From (12) and (13), we have

$$\sum_{s''_{-i}\in S_{-i}} U_i(x(s''_{-i},s_i),t_i(s''_{-i},s_i),s_i)p_i(s''_{-i})$$
$$> \sum_{s''_{-i}\in S_{-i}} U_i(x(s''_{-i},\phi_i(s_i)),t_i(s''_{-i},\phi_i(s_i)),s_i)p_i(s''_{-i}), \tag{14}$$

which is inequality (6) of Condition S. □

Remarks, The essential part of the proof of Proposition 1 is (9). Notice that the proof of the first equality of (9) depends on the assumption of private vales. It is easy to verify that in common values case the first equality of (9) does not generally hold. In addition, without the assumption of independence of prior probability the third equality of (9) does not hold. Next, for the number of agents (n) Dutta and Sen (1994a) extended their general characterization of Bayesian implementable social choice correspondences when $n \geq 3$ to the $n = 2$ case for economic environments. The environments will be satisfied if there is a transferable private good in which the utilities of both individuals are strictly increasing.[9] Thus, assuming transfer payments in this paper, Proposition 1 holds for $n = 2$ case.

An intuitive explanation of Proposition 1 is as follows. As mentioned in Palfrey and Srivastava (1993), in the environments satisfying the condition of NCD we can eliminate equilibria using simple and natural mechanisms, by specifying small side payments, or fines and rewards, that guarantee the implementation of an incentive compatible social choice function. Consider the example. There are two agents, each of whom has two types. That is, $N = \{1,2\}$ and $S_i = \{s_H, s_L\}$ for $i = 1, 2$. Suppose that the agents' types are independently distributed as follows:

$$p_i(s_H) = q \quad \text{and} \quad p_i(s_L) = 1 - q \quad \text{for} \quad i = 1,2 .$$

Suppose that $q \neq 1/2$. It is easy to check that the condition of NCD is satisfied since $q \neq 1/2$. Assume both agents are using strategy of always lying. Then either agent, say agent 2, is reporting s_H with probability $1-q$ and s_L with probability q. With truth telling, these probabilities are reversed. Without loss of generality, suppose $q > 1/2$. Now consider a transfer rule μ in addition to the allocation rule (x,t), where the rule calls for agent 2 to pay \$1 if agent 1 reports H, and for agent 1 to pay agent 2 $[\frac{q}{1-q} - \varepsilon]$ if agent 1 reports L. Such an 'augmented' allocation has the property that, for small ε, it makes agent 2 better off than (x,t) if agent 1 is always lying. The above claim is true regardless of agent 2's true type, since the transfer is constructed independent of his type. Thus, we can augment the direct mechanism to eliminate undesirable equilibria.

Furthermore, WSSS condition ensures that, even though consistent deception is used, Bayesian monotonicity holds. To show this, we consider the following social choice function (x', t') such that, irrespective of agent i's report, the social choice function $(x(\cdot, s_i), t(\cdot, s_i))$ is implemented. This selectively eliminates the potential equilibrium when the agents use consistent deception. The reason why this holds is as follows. For any agent and his or her reported type, the expected utility which he

[9] Refer to Palfrey (2002).

or she obtains when all other agents tell the truth is equal to that when all other agents use consistent deception. Combined the above fact with WSSS condition, there exists an agent i who uses the deception $\phi_i(s_i) \neq s_i$ and strictly prefers reporting his or her type truthfully to misrepresenting it. Using this fact, we can construct the augmented mechanism in which the deception is no longer an equilibrium.

In addition, notice that Proposition 1 implies that, under NCD, SS is a sufficient condition for the Bayesian implementation in our environment with side payments. In a more general environment, Matsushima (1993) showed that under SSS Condition 1 is equivalent to Bayesian monotonicity. Now, in Proposition 2 we provide the logical relation between the WSSS condition and Condition 1 in our environment .

Remark, Note that WSSS condition already incorporates the self-selection condition in its definition and it implies Bayesian monotonicity by Proposition 1. Thus, it is a sufficient condition for Bayesian implementation in our environment under quasi-linear utility function with side payments. A proof of the sufficient condition requires construction of a general mechanism as in Palfrey and Srivastava (1993).

Proposition 2. *Suppose that the agents' types are independently distributed. Moreover, suppose that a social choice function satisfies SS. The social choice function satisfies Condition 1 if it satisfies Condition S.*

Proof. From Condition S, if a social choice function (x,t) satisfies inequality (1) for a consistent deception ϕ, (6) holds for some $i \in N$ and $s_i \in S_i$. Fix $i \in N$ and $s_i \in S_i$. Consider the social choice function (x',t') such that, irrespective of the agent i's report, given a state of $s_i \in S_i$, the social choice function $(x',t')(s') = (x,t)(s'_{-i}, s_i)$ for all $s' \in S$.

At first, consider the lower part of inequality (7). By definition of (x',t') we have

$$U_i(x'(s'_{-i},\phi_i(s_i)), t'_i(s'_{-i},\phi_i(s_i)), \phi_i(s_i)) = U_i(x(s'_{-i},s_i)), t_i(s'_{-i},s_i), \phi_i(s_i)).$$

Thus the lower part of inequality (7) can be written as follows:

$$\sum_{s'_{-i}\in S_{-i}} [U_i(x(s'_{-i},s_i), t_i(s'_{-i},s_i)), \phi_i(s_i)) \\ -U_i(x(s'_{-i},\phi_i(s_i)), t_i(s'_{-i},\phi_i(s_i)), \phi_i(s_i))] p_i(s'_{-i}) \leq 0 . \tag{15}$$

The inequality follows from SS.

Now consider the upper part of inequality (7). From (9), we have the following:

$$\sum_{s_{-i}\in S_{-i}} [U_i(x'(\phi(s)),t_i'(\phi(s)),s_i)-U_i(x(\phi(s)),t_i(\phi(s)),s_i)]p_i(s_{-i})$$
$$=\sum_{s'_{-i}\in S_{-i}} U_i(x(s'_{-i},s_i),t_i(s'_{-i},s_i),s_i)-U_i(x(s'_{-i},\phi_i(s_i)),t_i(s'_{-i},\phi_i(s_i)),s_i)]p_i(s'_{-i})`>0\,.$$

Together the above inequality with (15), we have inequality (7) of Condition 1. □

IV. Discussions

In contrast to our paper, Palfrey and Srivastava (1993) showed that if types are independent, NCD is satisfied, and there is a money good, then a social choice function is incentive compatible if and only if it is Bayesian implementable. Notice that this result does not assume private values environments. Given the Palfrey and Srivastava's result, the issue here is whether or not NCD is satisfied in independent private values environments. If NCD is satisfied in the environments, conditions other than incentive compatibility are vacuous for Bayesian implementation. Therefore, it is important to provide an example of social choice function that satisfies incentive compatibility and Condition S, while NCD is not satisfied. Here we present two examples of the social choice function (x,t) that satisfies WSSS condition in case NCD does not hold.

Example 1. There are two agents, each of whom has two types. That is, $N=\{1,2\}$ and $S_i=\{s_H,s_L\}$ for $i=1,2$. There are four feasible allocations: a_{HH},a_{HL}, a_{LH},a_{LL}, where $a_{ij}\neq a_{kl}$ except for $i=k\in\{H,L\}$ and $j=l\in\{H,L\}$. The social choice function (x,t) is given by Table 1 below:

[Table 1] Social choice function for Example 1

		Agent 2	
		s_H	s_L
Agent 1	s_H	a_{HH}	a_{HL}
	s_L	a_{LH}	a_{HH}

Suppose that the agents' types are independently distributed as follows:

$$p_i(s_j \mid s_k) = p_i(s_j) = \frac{1}{2} \quad \text{for} \quad i = 1,2 \quad \text{and} \quad j,k \in \{H,L\}.$$

We can consider three cases for the possible forms of consistent deceptions:

First case, consider the following (pure) deception:
$\phi_1(s_H) = s_L$, $\phi_1(s_L) = s_L$ and $\phi_2 = id_2$, where id_i is the identity function, i.e., $id_i(s_i) = s_i$ for all i and $s_i \in S_i$. This case means that agent 1 uses the consistent deception but agent 2 tells the truth.

Clearly, $p_i(s_{-i} \mid s_i) = \hat{p}_i(s_{-i} \mid s_i, \phi_{-i}) = \frac{1}{2}$ for $i = 1,2$.
Thus, this deception ϕ is consistent.

Second case, consider the following deception:
$\phi_2(s_H) = s_L$, $\phi_2(s_L) = s_L$ and $\phi_1 = id_1$. This case means that agent 2 uses the consistent deception but agent 1 tells the truth.
Similarly to the first case, it is easy to verify that the deception ϕ is consistent.

Third case, $\phi_1(s_H) = s_L$, $\phi_1(s_L) = s_H$ and $\phi_2(s_H) = s_L$, $\phi_2(s_L) = s_H$.
Similarly to the first case, it is easy to verify that the deception ϕ is consistent.

In the above, we consider possible deceptions except pooling deceptions where some agents with different signals choose the same message, for example, $\phi_i(s_i) = s_L$ for all $s_i \in S_i$. The reason is that when prior has full support, such pooling deceptions cannot be consistent deceptions.

Now we consider the requirements to satisfy Condition S. For the first case, it suffices to check Condition S for the consistent deception ϕ to satisfy inequality (1). In state (s_H, s_H), we have

$$a_{HH} = (x,t)(s_H, s_H) \neq (x,t)(\phi_1(s_H), \phi_2(s_H)) = (x,t)(s_L, s_H) = a_{LH}.$$

Thus Condition S requires that inequality (6) be satisfied for agent 1 under the signal s_H, where $\phi_1(s_H) = s_L (\neq s_H)$. This condition is given by

$$\frac{1}{2}U_1(a_{HH}, s_H) + \frac{1}{2}U_1(a_{HL}, s_H) > \frac{1}{2}U_1(a_{LH}, s_H) + \frac{1}{2}U_1(a_{LL}, s_H). \tag{16}$$

In state (s_H, s_L), similarly to the above state it is easy to show that Condition S requires that inequality (16) hold.
In state (s_L, s_H), we have

$$a_{LH}=(x,t)(s_L,s_H)\neq(x,t)(\phi_1(s_L),\phi_2(s_H))=(x,t)(s_H,s_H)=a_{HH}\,.$$

Thus Condition S requires that inequality (6) be satisfied for agent 1 under the signal s_L, where $\phi_1(s_L)=s_H(\neq s_L)$. This condition is given by

$$\frac{1}{2}U_1(a_{LH},s_L)+\frac{1}{2}U_1(a_{LL},s_L)>\frac{1}{2}U_1(a_{HH},s_L)+\frac{1}{2}U_1(a_{HL},s_L)\,. \tag{17}$$

In state (s_L,s_L), similarly to the above state it is easy to show that Condition S requires that inequality (17) hold.

For the second case, it suffices to check Condition S for the consistent deception ϕ to satisfy inequality (1). Similarly to the above case, those conditions are given by

$$\frac{1}{2}U_2(a_{HH},s_H)+\frac{1}{2}U_2(a_{LH},s_H)>\frac{1}{2}U_2(a_{HL},s_H)+\frac{1}{2}U_2(a_{LL},s_H)\,, \tag{18}$$

$$\frac{1}{2}U_2(a_{HL},s_L)+\frac{1}{2}U_2(a_{LL},s_L)>\frac{1}{2}U_2(a_{HH},s_L)+\frac{1}{2}U_2(a_{LH},s_L)\,. \tag{19}$$

In third case, since all agents use the consistent deception, by definition Condition S requires that either (16)∼(17) for agent 1 or (18)∼(19) for agent 2 be satisfied.

From the results for three cases, Condition S requires that (16)∼(17) for agent 1 and (18)∼(19) for agent 2 be satisfied. Therefore, in the above example we know that the combination of SS and Condition S (WSSS condition) is equivalent to SSS condition.

Now we present an example in which the WSSS condition is weaker than SSS condition.

Example 2. There are two agents who have three types, respectively. That is, $N=\{1,2\}$ and $S_i=\{s_H,s_M,s_L\}$ for $i=1,2$. There are nine possible states: (s_H,s_H), (s_H,s_M), (s_H,s_L), (s_M,s_H), (s_M,s_M), (s_M,s_L), (s_L,s_H), (s_L,s_M), (s_L,s_L). The above states correspond to the following social choice function: $x(s_i,s_j)=a_{ij}$, where $a_{ij}\neq a_{kl}$ except for $i=k\in\{H,M,L\}$ and $j=l\in\{H,M,L\}$. Suppose that the agents' types are independently distributed as follows:

$$p_i(s_H)=p_i(s_L)=1/4 \quad \text{and} \quad p_i(s_M)=1/2 \quad \text{for} \quad i=1,2\,.$$

As in the example 1, pooling deceptions cannot be consistent deception.

In [Table 2], each row corresponds to a deception function ϕ_i. And each cell in the row represents the probability of observing a type s_i given a specific deception. For example, given deception function D2 the probability of observing a type s_L is 1/2. The reason is that using D2 agent i reports s_L when his or her type is s_M. The deception D1 is a truthful revelation, i.e., $\phi(s)=s$. It is easy to check that the probability of observing a type given deception D6 is equal to that given deception D1. Therefore, the only deception D6 other than truthful revelation is consistent. Now we find additional requirements from Condition S. It suffices to check the condition for the above consistent deception D6 satisfying inequality (1).

[Table 2] The probability of observing types given the deception function

Deception function		s_H	s_M	s_L
D1	$\phi_i(s_H)=s_H$, $\phi_i(s_M)=s_M$, $\phi_i(s_L)=s_L$	1/4	1/2	1/4
D2	$\phi_i(s_H)=s_H$, $\phi_i(s_M)=s_L$, $\phi_i(s_L)=s_M$	1/4	1/4	1/2
D3	$\phi_i(s_H)=s_M$, $\phi_i(s_M)=s_H$, $\phi_i(s_L)=s_L$	1/2	1/4	1/4
D4	$\phi_i(s_H)=s_M$, $\phi_i(s_M)=s_L$, $\phi_i(s_L)=s_H$	1/4	1/4	1/2
D5	$\phi_i(s_H)=s_L$, $\phi_i(s_M)=s_H$, $\phi_i(s_L)=s_M$	1/2	1/4	1/4
D6	$\phi_i(s_H)=s_L$, $\phi_i(s_M)=s_M$, $\phi_i(s_L)=s_H$	1/4	1/2	1/4

As in the above example we can consider three cases for the possible forms of consistent deceptions. For instance we have the deception for the first case

$$\phi_1(s_H)=s_L,\ \ \phi_1(s_M)=s_M,\ \ \phi_1(s_L)=s_L \ \text{ and } \ \phi_2=id_2 .$$

For brevity, let $(x,t)\tilde{P}^i(s_i)(x,t)(\cdot\,|\,\phi_i(s_i))$ denote

$$\sum_{s_{-i}\in S_{-i}} U_i(x(s),t_i(s),s_i)p_i(s_{-i}) > \sum_{s_{-i}\in S_{-i}} U_i(x(s_{-i},\phi_i(s_i)),t_i(s_{-i},\phi_i(s_i)),s_i)p_i(s_{-i}) .$$

From the results for three cases, it is easy to check that condition S is satisfied if the following conditions hold:

$$(x,t)\tilde{P}^1(s_H)(x,t)(\cdot\,|\,s_L),\ \ (x,t)\tilde{P}^2(s_H)(x,t)(\cdot\,|\,s_L),$$
$$(x,t)\tilde{P}^1(s_L)(x,t)(\cdot\,|\,s_H),\ \ (x,t)\tilde{P}^2(s_L)(x,t)(\cdot\,|\,s_H). \tag{20}$$

However, in the example the SSS condition states requires that

$$(x,t)\tilde{P}^i(s_j)(x,\,t)(\cdot\,|\,s_k),\ \text{where}\ \ i=1,2 \ \text{ and } \ j\neq k=H,L,M . \tag{21}$$

Therefore, in the above example the social choice function satisfies WSSS but does not satisfy SSS when NCD does not hold. The reason is that (20) implies (21) but the inverse does not hold.

Throughout this paper, we have restricted attention to the case where values are private. Whether or not the results in this paper are robust with respect to common values remains to be investigated. Moreover, our results consider pure deceptions (strategies). Whether the results can be extended to mixed Bayesian implementation using mixed deceptions (strategies) is an open question.

References

Dutta, B. and A. Sen (1994a), “2-person Bayesian implementation,” *Review of Economic design*, Vol. 1, pp. 41-54.

Dutta, B. and A. Sen (1994b), “Bayesian Implementation: The Necessity of Infinite Mechanisms,” *Journal of Economic Theory*, Vol. 64, pp. 130-141.

Jackson, M. (1991), “Full Bayesian Implementation,” *Econometrica*, Vol. 59, pp. 461-478.

Maskin, E. and T. Sjostrom (2002), Implementation theory: in: K. Arrow, A. Sen, and K. Suzumura (ed.), *Handbook of Social Choice and Welfare*, Vol. 1 (Elsevier Science, Amsterdam) , pp. 237-88.

Matsushima, H. (1993), “Bayesian Monotonicity with Side Payments,” *Journal of Economic Theory*, Vol. 59, pp. 107-121.

Mookherjee, D and S. Reichelstein (1990), “Implementation via Augmented Revelation Mechanisms,” *Review of Economic Studies*, Vol. 57, pp. 453-475.

Palfrey, T. (2002), Implementation Theory, in: K. Arrow, R. Aumann, and S. Hart (ed.), *Handbook of Game Theory with Economic Applications*, Vol. 3 (Elsevier Science, Amsterdam), p. 2310.

Palfrey, T and S. Srivastava (1989a), “Implementation with Incomplete Information in Exchange Economies,” *Econometrica*, Vol. 57, pp. 115-134.

Palfrey, T and S. Srivastava (1989b), “Mechanism Design with Incomplete Information: A Solution to the Implementation Problem,” *Journal of Political Economy*, Vol. 97, pp. 668-691.

Palfrey, T. (1992), “Implementation in Bayesian Equilibrium: the Multiple Equilibrium Problem in Mechanism Design,” in *Advances in Economic Theory: Sixth World Congress*, Vol. 1, J. J. Laffont, ed., Cambridge University Press, pp. 283-323.

Palfrey, R and S. Srivastava (1993), Bayesian Implementation, in *Fundamentals of Pure and Applied Economics,* 53 (Harwood Academic Publishers, New York).