計量經濟學報, 제16권 제1호(2005. 3. 31.), 33-52

Implementation of Allocation Rule with Strict Self-Selection*

Bong-Ju Kim**

This paper deals with the Bayesian implementation of an allocation rule in an environment with quasi-linear preferences. We analyze the implementation under the assumption that the principal does not know the prior distribution of agents' types, though the distribution is common knowledge among agents. Assuming that the agents' types are independent prior distributions, we show that an allocation rule with SSS can be fully implemented in perfect Bayesian equilibrium. However, in the case of dependent prior distribution of agents' types, we show that this result cannot hold. Hence, we suggest the sufficient condition for implementation.

Keywords: Implementation, Allocation Rule, Strict Self-selection
JEL classification: D78, D82

I. Introduction

The implementation problem is to design a mechanism achieving desirable outcomes according to a given allocation rule, across a domain of possible agent preferences. In order to implement the allocation rule, it is necessary that an equilibrium of mechanism produces desirable outcomes. However, the existence of the desired equilibrium depends on the equilibrium concept. For example, in the allocation problems of public goods, Green and Laffont(1977)

* 본 논문에 대해 귀중한 조언과 유익한 논평을 해주신 신성휘교수님과 익명의 심사위원들께 깊이 감사드립니다.
** KT 마케팅연구소 수석연구원, Tel: 031-727-6690, E-mail: kbongju@paran.com
투고일: 2004. 11. 11 심사일: 2004. 11. 15 최종심사완료일: 2004. 12. 29

and Walker(1980) showed that any ex post efficient, dominant strategy mechanism fails to balance the budget. D'Aspremont and Gerard-Varet(1979), and independently Arrow(1979), however, showed that budget balance could be overcome in a setting where agents have probabilistic beliefs over the types of other agents. The explanation for this positive result is as follows. The former mechanism requires dominant strategy incentive compatibility, i.e., truthful revelation of preference is the best response regardless of the potential announcements of the others. But this requirement is a very strong condition. Therefore, the latter assumes a weaker form of incentive compatibility, i.e., Bayesian incentive compatibility, and obtains a positive result. Thus, the Bayesian mechanism allows many allocation problems to be implemented. However, most Bayesian mechanisms have three major shortcomings. First, they rely heavily on Bayesian characteristics of agents, especially on the prior distribution of agents, e.g., the optimal auction mechanisms of Myerson(1981) and the bilateral trading mechanisms of Myerson and Satterthwaite(1983). Therefore, Bayesian mechanisms not only require that the prior beliefs of each agent be known to all agents, they also assume that the planner knows the agents' prior beliefs. Notably, it has been argued that this is too demanding in terms of informational requirements for the planner. In this paper we look at Bayesian mechanisms where the designer does not have to know the statistical information about the distribution of agents' characteristics.

The second problem with most Bayesian mechanisms, which is particularly prevalent in the incentive compatibility literature, is that they may possess multiple equilibria. Some of the equilibria may generate undesirable outcomes. Our paper eliminates the undesirable equilibria in Bayesian mechanisms by equilibrium refinement. Like Brusco(1998), who eliminated the undesirable equilibria in Bayesian mechanisms, we refine the Bayesian equilibrium to a perfect Bayesian equilibrium. While Brusco(1998) was successful in eliminating the undesirable equilibria, the mechanisms still depend on the priors of the agents. The last problem is that most Bayesian mechanisms deal with the case where agents are unable to form a coalition to collectively manipulate the decision rule. By focusing on the role of individual incentive constraints, they neglect the possibility of collusive agreements. The purpose of this paper is to present a mechanism that is immune to the aforementioned first and second problems. However, we do not address the third problem in this research. Laffont and Martimort(2000) and Che and Kim(2004) consider optimal mechanisms when agents can enter collusive agreements.

Contrary to the existing literature, we analyze an implementation under the assumption that the principal does not know the prior distribution of agents' types, though the distribution is common knowledge among the agents. While the present study is by no means the first attempt to resolve these two problems, the previous papers only treated specific problems. For example, Choi and Kim(1999) consider the problem of deciding whether to undertake a fixed-size public project. In contrast, the present study considers the general allocation rules.

For similar environments of this paper, Duggan(1998) showed that the principal virtually implements the principal's desired(second best) allocation rule. Contrary to this, however, we find sufficient conditions in which the principal fully implements the allocation rule. Specifically, we investigate the full implementation of an allocation rule with strict self-selection(SSS). We consider a mechanism with an augmented message space that includes the prior distribution, and we allow the agents to challenge each other's reports. Assuming the independent agents' prior distribution, we show that given any prior distribution an allocation rule with SSS can be fully implemented in perfect Bayesian equilibrium, even if the principal does not know the agents' prior distribution. However, under the assumption that the dependent prior distribution of agents' types, we show that this result cannot hold. Thus, we suggest the sufficient condition for the implementation. The remainderof the paper is organized as follows. In Section 2, we introduce the model and some notations. In Section 3, we present our mechanism and main results. Finally, we give our conclusions in Section 4.

Ⅱ. The model

We employ the model in which agents are incompletely informed about the preferences of other agents. Our model includes one principal and three or more agents. Let $N=\{1,\cdots,n\}$ denote a finite set of agents. Θ_i denotes the set of possible types for agent i with the generic element θ_i. A type for agent i, θ_i, specifies the preference and also his information about other agents. Let $\Theta = \times\Theta_i$ with the generic element $\theta = (\theta_i, \cdots, \theta_n)$. Assume that each Θ_i has finite elements, that is, $\Theta_i = \{\theta_{1i}, \cdots, \theta_{si}\}$.

Let P be a possible set of prior distributions of agents' types. We denote its generic element as p. The set P always includes a true prior distribution. If the prior distribution is common knowledge among the principal and agents, P is the singleton of a true prior distribution. In the interim stage, the agents know their individual type, and the distribution of other agents' types. Let $p_i(\theta_{-i} \mid \theta_i)$ denote agent i's conditional probability that other agents receive the profile θ_{-i} when agent i receives the signal θ_i. We assume that prior distributions are diffuse, i.e., for all $i \in N$, all $\theta \in \Theta$, and all $p \in P$, $p(\theta) > 0$. Furthermore, we assume that the principal does not know the prior distribution of agents' types though the distribution is common knowledge among the agents.

In the economy, m-goods and transferable money exist. Let $E = P \times \Theta$ denote the economic environment, and $e = (p, \theta)$ denote its generic element. A set of all feasible allocations is denoted by A. The allocation rule $x: E \rightarrow A$ is a function from state to allocation. Let $x(e) = (q_1(e), \cdots, q_n(e), t_1(e), \cdots, t_n(e))$, where $q_i: e \rightarrow R^m$ is the allocation of m-goods to agent i and $t_i : e \rightarrow R$ is the allocation of money to agent i.[2] Given $x(p, \theta)$, the utility function of agent i, who has type θ_i, is quasi-linear as follows:

$$U_i(x(p, \theta), \theta_i) = V_i(x(p, \theta), \theta_i) + t_i(p, \theta)$$

A mechanism is a pair (M, g), where $M = M^1 \times \cdots \times M^n$ with generic element

---

2) We assume that the allocation of m-goods depends on prior distribution. In an optimal auction and in a firm's pricing problems, the principal maximizes the expected payoff. Consequently, the allocation of m-goods depends on the prior distribution.

m and $g: M \rightarrow A$. The term M^i is the message space of agent i, while g is the outcome function.

Definition 1: x can be fully implemented if for all $e \in E$, $g(m(e)) = x(e)$.

Defining a direct mechanism, we consider the case in which prior distribution is common knowledge among the principal and agents. Knowing the prior distribution, the principal does not need the agents to report their prior distribution. In this environment, if $M^i = \Theta_i$ for all i, then (M, g) is a direct mechanism. In a direct mechanism, given allocation rule x, true distribution p, and θ_i, the expected utility to agent i reporting θ_i' is denoted by

$$EU_i(x(p, \theta_{-i}, \theta_i') \mid p(\theta_{-i} \mid \theta_i)) = \sum_{\theta_{-i}} U_i(x(p, \theta_{-i}, \theta_i'), \theta_i) p(\theta_{-i} \mid \theta_i)$$

Definition 2(SSS): In a direct mechanism, the allocation rule $x(p, \theta)$ satisfies SSS if for all $i \in N$, $\theta_i \in \Theta_i$ and $\theta_i'(\neq \theta_i) \in \Theta_i$,

$$EU_i(x(p, \theta_{-i}, \theta_i) \mid p(\theta_{-i} \mid \theta_i)) > EU_i(x(\theta_{-i}, \theta_i', p) \mid p(\theta_{-i} \mid \theta_i)).$$

Let $\alpha_i: \Theta_i \rightarrow \Delta\Theta_i$ denote a deception for agent i, where $\Delta\Theta_i = \{(\alpha_{1i}, \cdots, \alpha_{si}) \in R^S \mid \alpha_{mi} \geq 0, \sum_{m=1}^{s} \alpha_{mi} = 1\}$ and α_{mi} denote the probability that agent i reports θ_{mi} when agent i's type is θ_i. Note that a truth-telling strategy can be viewed as a special case of a deception in which the probability distribution over the Θ_i is degenerate that is, for all i and m, $\alpha_i(\theta_{mi}) = (0, \cdots, \overset{m-th}{1}, \cdots, 0)$. We denote $\alpha = (\alpha_1, \cdots, \alpha_n)$.

We define $p_i(\alpha_{-i}^{-1}(\theta_{-i}) \mid \theta_i) = p_i(m_{-i} = \theta_{-i} \mid \alpha_{-i}, \theta_i)$, which is the probability of observing θ_{-i} given agent i's type θ_i when the other agents use the deception α_{-i}.

Matushima(1993) defined consistent deception and no consistent deception as follows:

Ⅲ. Results

Consider the mechanism (M, g). It consists of a two-stage mechanism described as follows: in the first stage, the agents report their individual prior distribution and the agents' reports are then announced; in the second stage, the agents report their types and challenge each other's reports.

Let the message space of agent i be $M_1^i \times M_2^i \times M_3^i$, where

$$M_1^i = P,$$
$$M_2^i = \Theta_i,$$
$$M_3^i = W_i \cup 0.$$

We denote the generic element by $m^i = (m_1^i, m_2^i, m_3^i)$.

For each i, we define[3)]

$$W_i = \left\{ w^i = (w_1, \cdots, w_n) : (\Theta_{-i}, m_1^i) \to R^n \;\middle|\; \sum_{k=1}^{n} w_k(\Theta_{-i}, \mid m_1^i) = 0, \forall \Theta_{-i} \text{ and} \right.$$
$$\left. \left| w_i(\Theta_{-i} \middle| m_1^i) \right| < \varepsilon_{p^*}, \sum_{\Theta_{-i}} p_i^*(\Theta_{-i} \mid m_1^i) w_i(\Theta_{-i} \mid m_1^i) < 0 \right\},$$

where

$$\varepsilon_{p^*} = \min_{\Theta_1 \Theta_1'} [E_{\Theta_{-i}} \{U_i(x(p^*, \Theta_{-i}, \Theta_i), \Theta_i) - U_i(x(p^*, \Theta_{-i}, \Theta_i'), \Theta_i) \mid p^*\}].$$

If at least (n-1) agents agree on the same prior distribution in the first stage, agent i uses the same prior distribution for p^*. Otherwise, agent i uses for p^* the prior distributions agent 1 reported in the first stage. In defining W_i, we make a distinction between action sets and feasible sets. Consequently, the side transfer function W_i, which does not satisfy this definition, is not a feasible strategy. Without loss of generality, we assume that agents always use feasible strategies.

We partition M into D_1, D_2, D_3, D_4 as follows:

3) Side transfer payments satisfy the budget balancing condition.

$D_1 = \{m \mid m_1^i = m_1^j, m_3^i = 0, \ \forall i, j \in N\}$

$D_2 = \{m \mid m_1^j \neq m_1^i, m_1^i = m_1^k, \ \forall i, k \neq j, \text{ and } m_3^i = 0, \forall i\}$

$D_3 = \{m \mid \exists \ k \ st. m_3^k \neq 0\}$, and

$D_4 = \left\{m \mid m \notin \bigcup_{i=1}^{3} D_i\right\}$

We define the outcome function g as follows:

If $m \in D_1 \cup D_2$, then $g(m) = x(m_1^i, m_2)$.

If $m \in D_3$, then $g(m) = x(m_1^1, m_2) + (0, m_3^i(m_2^{-i}))$, where $i = \min\{k \mid m_3^k \neq 0\}$ and $m_2^{-i} = (m_2^1, \cdots, m_2^{i-1}, m_2^{i+1}, \cdots, m_2^n)$.

If $m \in D_4$, then $g(m) = x(m_1^1, m_2)$.

The underlying idea of the mechanism is rather simple. Each agent is asked to report the prior distribution at the first stage. Next the agents report their individual type and whether they believe that everybody else is reporting honestly at the second stage. The agent's third message represents the latter. Zero (0) means that an agent believes that everybody else is reporting honestly. However, non-zero transfers mean that an agent believes that someone else is reporting dishonestly. If everybody claims the agents are reporting honestly then the allocation rule is implemented. Side transfers are designed so that whenever the other agents tell the truth the deviating agent obtains a negative transfer and the deviation is unprofitable. Truth-telling is therefore an equilibrium. Whenever a deception is used, it is not equilibriumfor all agents to agree with the deception. In this case, at least one agent obtains gains for proposing a non-zero transfer. In the parlance of Mookerjee and Reichelstein (1990), deception is selectively eliminated from being an equilibrium if an agent is allowed to propose a non-zero transfer. Finally, the mechanism has been constructed to avoid the addition of any new equilibrium involving a non-zero transfer. No such equilibrium can arise for any proposal with a non-zero transfer. In this case, the agent has no best response.[4)]

4) There is therefore an "open set"problem, and no best response exists for deviating agent.

For convenience of exposition, we denote the second message of agent i, i.e., m_2^i, by a deception function α_i. Let $\alpha_i : \Theta_i \rightarrow \Delta\Theta_i$ denote a deception for agent i where $\Delta\Theta_i = \left\{(\alpha_{1i}, \cdots, \alpha_{si}) \in R^S \mid \alpha_{mi} \geq 0, \sum_{k=1}^{s} \alpha_{ki} = 1\right\}$ and α_{ki} denote the probability that agent i reports α_{ki} when agent i's type is Θ_{mi}. Note that a truth-telling strategy can be viewed as a special case of a deception in which the probability distribution over the Θ_i is degenerate that is, for i and Θ_{mi}, $\alpha_i(\Theta_{mi}) = (0, \cdots, \overset{m-th}{1}, \cdots, 0)$. In this case, we denote $m_2^i = \Theta_i$.

We denote $\alpha = (\alpha_1, \cdots, \alpha_n)$. We define $p(\alpha_{-i}^{-1}(\Theta_{-i}) \mid \Theta_i) = p(m_2^{-i} = \Theta_{-i} \mid \alpha_{-i}, \Theta_i)$, which is the probability of observing Θ_{-i} given agent i's type Θ_i when the other agents use the deception α_{-i}.

We analyze an implementation with the assumption that the principal does not know the prior distribution of agents' types though the distribution is common knowledge among the agents. First, we assume that the agents' types are independent. Then, in proposition 1, we show that the allocation rule with SSS can be fully implemented in a perfect Bayesian equilibrium, even if the principal does not know the agents' prior distribution. Next, in proposition 2, we suggest an additional sufficient condition for the implementation if the agents' types distribution are dependent.

Now we prove the proposition 1, which uses the above mechanism (M, g). However, the side transfer functions are revised as follows:

$$W_i = \left\{ w : \Theta_{-i} \rightarrow R^n \;\middle|\; \begin{array}{l} \sum_{k=1}^{n} w_k(\Theta_{-i}) = 0, \forall \Theta_{-i} \mid w_i(\Theta_{-i}) \mid < \varepsilon \\ \text{and } \sum_{\Theta_{-i}} p_i^*(\Theta_{-i}) w_i(\Theta_{-i}) < 0 \end{array} \right\},$$

where ε takes a positive small number.

Remark: Contrary to the dependent distributions, conditional distributions do not depend on agents' types. Thus the side transfer functions do not depend on agents' types reporting, and consequently the agents can not improve the payoff by manipulating the types' distribution by way of reporting a false type.

Proposition 1. Suppose that $n \geq 3$, agents have an independent distribution, and Θ is finite. Let x be an allocation rule that satisfies SSS. Then x can be fully implemented in a perfect Bayesian equilibrium, even if the principal does not know the agents' prior distribution.

Proof. Take any e. The proof consists of the following two steps:

Step 1: We show that given $e=(p,\Theta)$, for every i, $m^i=(p,\Theta_i,0)$ is a Bayesian equilibrium. Now we show that $m_1^i=p$ is a best response. In the first stage, if everyone other than i reports p, then i cannot alter the allocaton rule $x(p,\ \cdot\ )$ by reporting $\hat{p}$. So, $m_1^i=p$ is a best response. Next we show that for all i, $m_2^i=\Theta_i$ and $m_3^i=0$ is a Bayesian equilibrium. We consider the following two cases. First, by definition of W_i, if $m_2^i=\Theta_i$, then the agent can suffer losses $\sum_{\Theta_{-i}} p_i(\Theta_{-i})w_i(\Theta_{-i})$ from reporting $m_3^i \neq 0$. Second, if $m_2^i \neq \Theta_i$, then agent i can losses from $m_2^i \neq \Theta_i$ in addition to reporting $m_3^i \neq 0$. Therefore, for all i, $m_2^i=\Theta_i$ and $m_3^i=0$ is a Bayesian equilibrium.

Step 2: We show that if any Bayesian equilibrium exists, the equilibrium outcome is identical with the allocation rule. It suffices to show that any Bayesian equilibrium should satisfy the following three conditions:

(i) In any perfect Bayesian equilibrium, $m_3^i=0$ for all agents.

Now we consider the second stage of the mechanism. Let p^* be the prior distribution adopted at the first stage. Suppose that $m_3^j \neq 0$ for some j and Θ_j. The agent with the lowest index is chosen as agent i. We assume that $\alpha_i(\Theta_i)$ is a mixed strategy of reporting Θ_{mi} with probability α_{mi} where $m=1,\cdots,s$. Without loss of generality, the elements of the set $\widehat{\Theta}_{-i}=\{\Theta_{1i},\cdots,\Theta_{wi}\}$ (where $w \leq s$) have a positive probability and satisfy the definition of W_i. Then, one of the following holds:

(C 1) $\quad \varepsilon > \sum_{l=1}^{w} \alpha_{li} \sum_{\Theta_{-i}} p_i(\alpha_{-i}^{-1}(\Theta_{-i}))w_i(\Theta_{-i}) > 0$,

(C 2) $\quad 0 > \sum_{l=1}^{w} \alpha_{li} \sum_{\Theta_{-i}} p_i(\alpha_{-i}^{-1}(\Theta_{-i}))w_i(\Theta_{-i}) > -\varepsilon$,

(C 3) $\quad \sum_{l=1}^{w} \alpha_{li} \sum_{\Theta_{-i}} p_i(\alpha_{-i}^{-1}(\Theta_{-i}))w_i(\Theta_{-i}) = 0.$

In (C 1), agent i can improve the payoff by reporting $kw_i(\cdot)$ as follows: Since Θ_{-i} is finite, k exists such that

$$\min_{\Theta_{-i}}[\varepsilon / \mid w_i(\Theta_{-i}) \mid] > k > 1.$$

Thus, agent i can improve the payoff by reporting $kw_i(\cdot)$. This outcome contradicts the statement that reporting $m_3^i \neq 0$ is a Bayesian equilibrium.

In (C 2), agent i can improve the payoff by reporting $kw_i(\cdot)$ such that $0 < k < 1$.

This result contradicts the statement that reporting $m_3^i \neq 0$ is a Bayesian equilibrium.

In the case of (C 3), we divide the following cases:

(C 3.1) $\quad p_i^*(\Theta_{-i}) = p_i(\alpha_{-i}^{-1}(\Theta_{-i}))$,

(C 3.2) $\quad p_i^*(\Theta_{-i}) \neq p_i(\alpha_{-i}^{-1}(\Theta_{-i}))$.

In the case of (C 3.1), inserting (C 3.1) into (C 3) implies

$$\sum_{l=1}^{m} \alpha_{li} \sum_{\Theta_{-i}} p_i^*(\Theta_{-i})w_i(\Theta_{-i}) = 0 \tag{1}$$

By the definition of W_i, the left-hand side of (1) has a negative value, which is a contradiction.

In the case of (C 3.2), (C 3) implies $\sum_{\Theta_{-i}} p_i^*(\alpha_{-i}^{-1}(\Theta_{-i}))w_i(\Theta_{-i}) = 0.$

But $p_i^*(\Theta_{-i}) \neq p_i(\alpha_{-i}^{-1}(\Theta_{-i}))$ implies that $\widehat{w}_i(\cdot)$ exists such that:

$$\sum_{\Theta_{-i}} p_i^*(\Theta_{-i})\widehat{w}_i(\Theta_{-i}) < 0,$$

$$\sum_{\Theta_{-i}} p_i(\Theta_{-i})\widehat{w}_i(\Theta_{-i}) > 0.$$

Thus, agent i can improve the payoff by reporting $\widehat{w}_i(\cdot)$ instead of $w_i(\cdot)$, which is a contradiction.

(ii) In any perfect Bayesian equilibrium, $m_2^i = \Theta_i$ for all agents i.

First, we show that in any Bayesian equilibrium, $p^*(\Theta_{-i}) = p(\alpha_{-i}^{-1}(\Theta_{-i}))$.

Suppose that $p^*(\Theta_{-i}) \neq p(\alpha_{-i}^{-1}(\Theta_{-i}))$ for some Bayesian equilibrium. As shown in (C 3.1), agent i can improve the payoff by reporting $\widehat{w}_i(\cdot)$ instead of $w_i(\cdot)$, which is a contradiction.

Next, suppose that $m_2^i \neq \Theta_i$. There exist $\widehat{\Theta}_i(\neq\Theta_i)$ which is a best response at Θ. Since $\widehat{\Theta}_i$ is a best response, this implies:

$$EU_i(x(p^*, \Theta_{-i}, \widehat{\Theta}_i)\Theta_i) \mid p(\alpha_{-i}^{-1}(\Theta_{-i})) \geq EU_i(x(p^*, \Theta_{-i}, \Theta_i)\Theta_i) \mid p(\alpha_{-i}^{-1}(\Theta_{-i}))$$

Since $p^*(\Theta_{-i}) = p(\alpha_{-i}^{-1}(\Theta_{-i}))$, this equation implies:

$$EU_i(x(p^*, \Theta_{-i}, \widehat{\Theta}_i)\Theta_i) \mid p^*(\Theta_{-i})) \geq EU_i(x(p^*, \Theta_{-i}, \Theta_i)\Theta_i) \mid p^*(\Theta_{-i}))$$

This result contradicts the assumption that the allocation rule satisfies SSS.

(iii) In any Bayesian equilibrium, the allocation rule adopted in the first period is $x(p, \cdot)$.

We consider the following cases, which alter the allocation rule.

First, in the first stage, at least (n-1) agents report the same false report p^* where $p \neq p^*$. By Lemma A1 of the appendix, there exist i and Θ_{-i} such that $p_i(\Theta_{-i}) \neq p_i^*(\Theta_{-i})$. Then agent i can improve his/her payoff by reporting $m_3^i \neq 0$, which contradicts (i). Second, two agents or more report the first message, which is different from the other agents' messages. In this case, agent

1's first message determines the allocation. Hence, if $m_1^i = p$, then the allocaton rule is $x(p, \Theta)$. But, if $m_1^i(=p^*) \neq p$, there exist i and Θ_{-i} such that $p_i(\Theta_{-i}) \neq p_i^*(\Theta_{-i})$. So agent i can improve his/her payoff by reporting $m_3^i \neq 0$, which contradicts (i).■

Proposition 1 can not extend the environments wherein the agents'types distribution are dependent. Hence, we suggest an additionalsufficient condition for the implementation as follows.

Definition 3(Condition S1): A possible set of prior distributions of agents' types, P, satisfies Condition S1 if the following information condition holds: For all pairs p, $p^*(\neq p) \in P$ and all $\overline{\Theta} \in \Theta$, there exist some i and $\Theta_{-i} \in \Theta_{-i}$ such that $p_i(\Theta_{-i} \mid \overline{\Theta}_i) \neq p_i^*(\Theta_{-i} \mid \overline{\Theta}_i)$.
Remark: By lemma A1 in the appendix, if any element of P is independent prior distribution, P satisfies Condition S1.

Proposition 2. Suppose that $n \geq 3$, agents have an dependent distribution, and Θ is finite. A possible set of prior distributions of agents' types, P, satisfies Condition S1. Let x be an allocation rule that satisfies SSS. Then x can be fully implemented in a perfect Bayesian equilibrium, even if the principal does not know the agents' prior distribution.

Proof. We prove the theorem that uses the above mechanism (M, g). The proof consists of following two steps:

Step 1: We show that given $e=(p, \Theta)$, for every i, $m^i = (p, \Theta_i, 0)$ is a Bayesian equilibrium. In the same way as the step 1 of proposition 1, we show that $m_1^i = p$ is a best response. Next we show that for all i, $m_2^i = \Theta_i$ and $m_3^i = 0$ is a Bayesian equilibrium. Suppose that it is not. Some agent i deviates from that. We consider the following two deviating cases. First, if $m_2^i = \Theta_{i'}$, by definition of W_i, she suffer the losses $\sum_{\Theta_{-i}} p_i(\Theta_{-i} \mid \Theta_i) w_i(\Theta_{-i} \mid \Theta_i)$ from reporting $m_3^i \neq 0$. So she does not deviate. Second, if $m_2^i \neq \Theta_i$, agent i can report $m_3^i = 0$ or $m_2^i \neq 0$. In the first case, for the allocation x satisfies SSS, agent i can not gain from reporting $m_1^i \neq \Theta_i$. In the second case, agent i has a positive expected transfer from reporting $m_2^i \neq 0$.[5)] However, W_i sets the ε_p^* to be small enough so that the gains from the deviation are less than the losses. For all i, $m_2^i = \Theta_i$ and $m_3^i = 0$ is a Bayesian equilibrium. Therefore, for every i, $m^i = (p, \Theta_i, 0)$ is a Bayesian equilibrium.

Step 2: We show that if any Bayesian equilibrium exists the equilibrium outcome is identical with the allocation rule. It suffices to show that any Bayesian equilibrium should satisfy the following three conditions:

(i) In any Bayesian equilibrium, $m_3^i = 0$ for every agent i.

Now we consider the second stage of mechanism. Let p^* be the prior distribution adopted at the first stage. Suppose that $m_3^j \neq 0$ for some j and Θ_j. The agent with the lowest index is chosen as agent i. We assume that $\alpha_i(\Theta_i)$ is a mixed strategy of reporting Θ_{mi} with probability Θ_{mi} where $m = 1, \cdots, s$. Without loss of generality, the elements of the set $\widehat{\Theta}_i = \{\Theta_{1i}, \cdots, \Theta_{wi}\}$(where $w \leq s$) have a positive probability and satisfy the definition of W_i. Then, one of the following holds:

(C 1) $$\varepsilon_{p^*} > \sum_{l=1}^{w} \alpha_{li} \sum_{\Theta_{-i}} p(\alpha_{-i}^{-1}(\Theta_{-i}) \mid \Theta_i^l) w_i(\Theta_{-i} \mid \Theta_i^l) > 0,$$

(C 2) $$0 > \sum_{l=1}^{w} \alpha_{li} \sum_{\Theta_{-i}} p(\alpha_{-i}^{-1}(\Theta_{-i}) \mid \Theta_i^l) w_i(\Theta_{-i} \mid \Theta_i^l) > -\varepsilon_{p^*},$$

(C 3) $$\sum_{l=1}^{w} \alpha_{li} \sum_{\Theta_{-i}} p(\alpha_{-i}^{-1}(\Theta_{-i}) \mid \Theta_i^l) w_i(\Theta_{-i} \mid \Theta_i^l) = 0.$$

5) Contrary to the independent distributions, conditional distributions depend on agents' types so that agents can not improve the payoff through manipulating types' distribution by way of reporting a false type. Thus we consider this deviation.

In (C 1), agent i can improve the payoff by reporting $kw_i(\cdot)$ as follows: Since Θ_{-i} is finite, k exists such that $\min_{\{\widehat{\Theta}_{-i},\, l=1,\cdots,w\}}[\varepsilon_{p^*}/\mid w_i(\Theta_{-i}\mid\Theta_i^l)\mid] > k > 1$.

Thus, agent i can improve the payoff by reporting $kw_i(\cdot)$. This outcome contradicts the statement that reporting $m_3^i \neq 0$ is a Bayesian equilibrium.

In (C 2), agent i can improve the payoff by reporting $kw_i(\cdot)$ such that $0 < k < 1$. This result contradicts the statement that reporting $m_3^i \neq 0$ is a Bayesian equilibrium.

In the third case, we divide the following cases:

(C 3.1) For all $\Theta_{ki} \in \widehat{\Theta}_i$, $p^*(\Theta_{-i}\mid\Theta_i^k) = p(\alpha_{-i}^{-1}(\Theta_{-i})\mid\Theta_i^k)$.

(C 3.2) For some $\Theta_{ki} \in \widehat{\Theta}_i$, $p^*(\Theta_{-i}\mid\Theta_i^k) \neq p(\alpha_{-i}^{-1}(\Theta_{-i})\mid\Theta_i^k)$

In the case of (C 3.1), inserting (C 3.1) into (C 3) implies

$$\sum_{l=1}^{w}\alpha_{li}\sum_{\Theta_{-i}}p_i^*(\Theta_{-i}\mid\Theta_{li})w_i(\Theta_{-i}\mid\Theta_{li}) = 0 \tag{2}$$

By the definition of W_i, the left-hand side of (2) has a negative value, which is a contradiction.

In the case of (C 3.2), we divide $\widehat{\Theta}_i$ into the following two cases:

(C 3.2.1) For $\Theta_{ki} \in \widehat{\Theta}_i$ such that $\sum_{\Theta_{-i}} p(\alpha_{-i}^{-1}(\Theta_{-i})\mid\Theta_i^k)w_i(\Theta_{-i}\mid\Theta_i^k) \neq 0$, agent i can improve his payoff in the same manner as (C 1) and (C 2).

(C 3.2.2) For $\Theta_{ki} \in \widehat{\Theta}_i$ such that, $\sum_{\Theta_{-i}} p(\alpha_{-i}^{-1}(\Theta_{-i})\mid\Theta_i^k)w_i(\Theta_{-i}\mid\Theta_i^k) = 0$, $p^*(\Theta_{-i}\mid\Theta_i^k) \neq p(\alpha_{-i}^{-1}(\Theta_{-i})\mid\Theta_i^k)$ implies that there exist $\widehat{w}_i(\cdot)$ such that:

$$\sum_{\Theta_{-i}} p^*(\Theta_{-i}\mid\Theta_i^k)\widehat{w}_i(\widehat{\Theta}_{-i}\mid\Theta_i^k) < 0,$$

$$\sum_{\Theta_{-i}} p(\alpha_{-i}^{-1}(\Theta_{-i})\mid\Theta_i^k)\widehat{w}_i(\Theta_{-i}\mid\Theta_i^k) > 0.$$

Thus, agent i can improve his payoff by reporting $\widehat{w}_i(\cdot)$ instead of $w_i(\cdot)$, which is a contradiction.

(ii) In any Bayesian equilibrium, $m_2^i = \Theta_i$ for every agent i.

In the same manner as (ii) of proposition 1, this can be easily demonstrated.

(iii) In any Bayesian equilibrium, the allocation rule adopted in the first stage is $x(p, \cdot)$.

By the assumption of condition S1, we can show this in the same manner as (iii) of proposition 1. Condition S1 is needed for the proof because, contrary to independent types for dependent types, a possible set of prior distributions of agents' types does not satisfy condition S1, as shown by lemma A2 in the appendix.■

Throughout this paper, we have confined our attention to allocation rules with SSS. In most studies on implementation, SSS is not a necessary condition for implementation whereas self-selection (Bayesian incentive compatibility) is required. In order to use Bayesian incentive compatibility as a necessary condition for implementation, we relax the implementability definition, and introduce a new condition as follows.

Definition 4(Condition S2): The information condition satisfies condition S2 if the following condition holds:

For every $i \in N, \Theta_i \in \Theta_i$, and $\Theta_i' \in \Theta_i/\{\Theta_i\}$, $p_i(\Theta_{-i} \mid \Theta_i) \neq p_i(\Theta_{-i} \mid \Theta_i')$ for some $\Theta_{-i} \in \Theta_{-i}$.

Under this assumption, according to Matushima(1993), any allocation rule with self-selection is close to another allocation rule with SSS. Close means that the allocation rule x is exactly implementable with a minor change of transfer.[6] If we combine this with the proposition 2, we obtain the following result:

Proposition 3: Suppose that $n \geq 3$. A possible set of prior distributions of

6) "Virtually implementable" means that the allocation rule x is implementable with a minor change of the allocation rule as well as transfers. "Exactly implementable" is much stronger than "virtually implementable."

agents' types, P, satisfies Condition S1 and S2. Then an allocation rule x with Bayesian incentive compatibility can be exactly implemented, even if the principal does not know the agents' prior distribution.

Ⅳ. Conclusion

This paper deals with the Bayesian implementation of an allocation rule in an environment with quasi-linear preferences. We analyze the implementation under the assumption that the principal does not know the prior distribution of agents' types, though the distribution is common knowledge among agents. Assuming that the agents' types are independent prior distributions, we show that an allocation rule with SSS can be fully implemented in perfect Bayesian equilibrium. However, in the case of a dependent prior distribution of agents' types, we show that this result cannot hold. Hence, we suggest the sufficient condition, i.e., condition S1, for the implementation. Duggan(1998) showed that the principal virtually implements her desired(second best) allocation rule. Contrary to this, however, we find sufficient condition in which the principal fully implements the allocation rule, even if she does not know the agents' prior distribution.

[References]

Arrow, K(1979), "The Property Rights Doctrine and Demand Revelation under Complete and Incomplete Information," *Economics and Human Welfare*, Academic Press, New York.

Brusco, S(1998), "Unique Implementation of the Full Surplus Extraction Outcome in Auction with Correlated Types," *Journal of Economic Theory*, 80, 185-200.

Choi, J. and T. Kim(1999), "A Nonparametric Efficient Public Good Decision Mechanism : Undominate Bayesian Implementation," *Games and Economic Behavior*, 27, 64-85.

Che, Y.-K. and J. Kim(2004), "Collusion-Proof Implementation of Optimal Mechanisms, mimeo, University of Wisconsin.

D'Aspremont, C., Gerard-Varet, L.A.(1979), "Incentives and Incomplete Information. *Journal of Public Economics*, 11, 25-45.

Duggan, J(1998), "An Extensive Form Solution to the Adverse Selection Problem in Principal/multi-agent Environments," *Review of Economic Design*, 3, 167-191.

Green, J. and J.-J. Laffont(1977), "Characterization of Satisfactory Mechanisms for the Revelation of the Preferences for Public Goods," *Econometrica*, 45, 427-438.

Jackson, M.O.(2000), "Mechanism Theory," Forthcoming in the Encyclopedia of Life Support Systems.

Laffont, J.-J. and D. Martimort(2000), "Mechanism Deign with Collusion and Correlation," *Econometrica*, 68, 309-342.

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

Mookherjee, D., Reichelstein, S.(1990), "Implementation via Augmented Revelation Mechanisms," *Review of Economic Studies*, 57, 453-475.

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

Myerson, M. and M. Satterthwaite(1983), "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29, 265-281.

Palfrey, R. and S. Srivastava(1989), "Mechanism Design with Incomplete Information: A Solution to the Implementation Problem," *Journal of Political Economy*, 97, 668-691.

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

Appendix

Lemma A1: Suppose that agents' types are independent. If there exits $\overline{\Theta}$ such that for all i and $\Theta_{-i}, p_i(\Theta_{-i} \mid \overline{\Theta}_i) = p_i^*(\Theta_{-i} \mid \overline{\Theta}_i)$, then for all $\Theta, p(\Theta) = p^*(\Theta)$.

Proof: Suppose there exists some Θ' such that $p(\Theta') \neq p^*(\Theta')$. By definition of the conditional probability, we have

$$p_i(\Theta'_{-i} \mid \Theta'_i) = \frac{p(\Theta')}{p_i(\Theta'_i)} \text{ and } p_i^*(\Theta'_{-i} \mid \Theta'_i) = \frac{p^*(\Theta')}{p_i^*(\Theta'_i)} \qquad \text{(A1)}$$

We choose agent $k(k \neq i)$. By the assumption of independent types, we have

$$p_k(\Theta_{-k} \mid \Theta_k) = p_k(\Theta_{-k}) = \frac{p(\Theta)}{p_k(\Theta_k)} = \prod_{j \neq k} p_j(\Theta_j)$$

Therefore $p_k(\Theta_{-k}) = \prod_{j \neq k} p_j(\Theta_j)$ for all Θ_{-k}. (A2)

In the same manner, $p_k^*(\Theta_{-k}) = \prod_{j \neq k} p_j^*(\Theta_j)$ for all Θ_{-k}. (A3)

By for all i and $\Theta_{-i}, p_i(\Theta_{-i} \mid \overline{\Theta}_i) = p_i^*(\Theta_{-i} \mid \overline{\Theta}_i)$ and the assumption of independent types, $p_k(\Theta_{-k}) = p_k^*(\Theta_{-k})$ for all Θ_{-k}.

By (A2) and (A3), we have $\sum_{\Theta_{-i-k}} \prod_{j \neq k} p_j(\Theta_j) = \sum_{\Theta_{-i-k}} \prod_{j \neq k} p_k^*(\Theta_j)$, where $\Theta_{-i-k} = \Theta_1 \times \cdots \times \Theta_{i-1} \times \Theta_{i+1} \times \cdots \times \Theta_{k-1} \times \Theta_{k+1} \times \cdots \times \Theta_n$.

Therefore, for all Θ_i, $p_i(\Theta_i) = p_i^*(\Theta_i)$. So (A1) and independent types implies $p(\Theta') = p^*(\Theta')$ which is a contradiction. ■

Lemma A2: Suppose that agents' types are correlated. Lemma A1 always does not hold.

Proof: To prove it, we suggest the following example. There are two agents who have two types respectively. Let $\Theta_i = \{\underline{\Theta}, \overline{\Theta}\}$ for $i = 1, 2$. There exits $\underline{\Theta}$ such that for all i and Θ_{-i}, $p_i(\Theta_{-i} \mid \underline{\Theta}) = p_i^*(\Theta_{-i} \mid \underline{\Theta})$, but $p(\Theta) \neq p^*(\Theta)$.

We consider the following two prior distributions.

Agent 2

Agent 1		$\Theta_2 = \underline{\Theta}$	$\Theta_2 = \overline{\Theta}$
	$\Theta_1 = \underline{\Theta}$	$\frac{3}{8}$	$\frac{2}{8}$
	$\Theta_1 = \overline{\Theta}$	$\frac{2}{8}$	$\frac{1}{8}$

<Figure A1> prior distribution $p(\Theta)$

Agent 2

Agent 1		$\Theta_2 = \underline{\Theta}$	$\Theta_2 = \overline{\Theta}$
	$\Theta_1 = \underline{\Theta}$	$\frac{2}{5}$	$\frac{4}{15}$
	$\Theta_1 = \overline{\Theta}$	$\frac{4}{15}$	$\frac{1}{15}$

<Figure A2> prior distribution $p^*(\Theta)$

From the $p(\Theta)$ and $p^*(\Theta)$, we have

$$p_1(\underline{\Theta} \mid \underline{\Theta}) = p_1^*(\underline{\Theta} \mid \underline{\Theta}) = \frac{3}{5}, \; p_1(\overline{\Theta} \mid \underline{\Theta}) = p_1^*(\overline{\Theta} \mid \underline{\Theta}) = \frac{2}{5};$$
$$p_2(\underline{\Theta} \mid \underline{\Theta}) = p_2^*(\underline{\Theta} \mid \underline{\Theta}) = \frac{3}{5}, \; p_2(\overline{\Theta} \mid \underline{\Theta}) = p_2^*(\overline{\Theta} \mid \underline{\Theta}) = \frac{2}{5}.$$

But it is obvious that $p(\Theta) \neq p^*(\Theta)$.■

[Abstract]

강한 자기선택을 지닌 배분규칙의 실행

김 봉 주

본 연구는 대리인들이 준선형 선호를 갖는 환경에서 배분규칙의 베이지안 실행 문제를 살펴본다. 사전확률분포가 대리인들간에는 공통지식(common knowledge)이나 주인은 이를 알지 못한다고 가정한 후 실행문제를 분석한다. 이때 대리인의 유형이 독립인 사전확률분포를 갖는 경우, 강한 자기선택(SSS)을 갖는 배분규칙은 완전 베이즈 균형(Perfect Bayesian Equilibrium)으로 실행할 수 있음을 보였다. 그러나 대리인들의 유형이 상관된 사전확률분포를 갖는 경우 이러한 결과는 유지될 수 없음을 보였다. 따라서 이러한 경우 실행을 위한 충분조건을 제시한다.

핵심용어: 실행, 배분규칙, 강한 자기선택.