Review of Economic Studies (1985) LII, 223-229 0034-6527/86/00150223$02.00

Implementation in Dominant Strategies under Complete and Incomplete Information

RAFAEL REPULLO
London School of Economics

This paper shows that if a social choice rule can be implemented in dominant strategies by an indirect mechanism, but there does not exist a direct mechanism that implements it in dominant strategies, then it must be the case that the original indirect mechanism does not implement the social choice rule in Nash strategies (under complete information) or in Bayesian strategies (under imcomplete information).

1. INTRODUCTION

A social choice rule is a correspondence that selects a set of optimal social states for each possible configuration of agents' characteristics. When these characteristics are not publicly known, it is assumed that the planner devises a mechanism, that is, a rule which specifies a social state for each vector of strategies chosen by the agents. A mechanism is said to implement a social choice rule in dominant strategies if (i) for every agent and every characteristic that he may have there exists at least one strategy which is dominant, and (ii) the social states that obtain when agents play these dominant strategies belong to the corresponding social choice set.

A special kind of mechanisms are those in which the strategy set for each agent coincides with the set of his possible characteristics. For these mechanisms, which are called direct, one may define an alternative notion of implementation: a direct mechanism truthfully implements a social choice rule in dominant strategies if (i) for every agent and every characteristic that he may have truth-telling is a dominant strategy, and (ii) the social state that obtains when agents truthfully report their characteristics belongs to the corresponding social choice set.

An important result in the literature on implementation, known as the revelation principle, establishes that for any mechanism that implements a social choice rule in dominant strategies there exists a direct mechanism which truthfully implements it in dominant strategies. This result, however, does not imply that when we consider implementation in dominant strategies there is no loss of generality in restricting attention to direct mechanisms, since there is nothing that guarantees that agents will always choose the truth-telling dominant strategies where alternative dominant strategies exist. Moreover, these alternative dominant strategies need not correspond to dominant strategies in the original mechanism, which suggests that the direct mechanism given by the revelation principle may not implement the given social choice rule in dominant strategies. The purpose of this paper is to investigate the reason why this may happen.

Traditionally, the analysis of implementation in dominant strategies has been developed without reference to the information that agents are assumed to have about each other's characteristics. This is not entirely satisfactory, because given a mechanism that implements a social choice rule in dominant strategies, there may be alternative Nash equilibria (under complete information) or Bayesian equilibria (under incomplete information) which cannot be ruled out a priori, and which may yield outcomes outside the social choice set. In this paper we show that this problem is responsible for the possible failure of the direct mechanism given by the revelation principle to implement the social choice rule in dominant strategies.

The paper is organized as follows. Section 2 presents an example to motivate the main results of the paper. Section 3 sets out the notation and basic definitions. Section 4 shows that if a social choice rule can be implemented in dominant strategies by a mechanism, but the direct mechanism given by the revelation principles does not implement it in dominant strategies, then it must be the case that the original mechanism does not implement the social choice rule in Nash strategies (under complete information) or in Bayesian strategies (under incomplete information). Finally, Section 5 contains a summary of the paper and some conclusions.

2. AN EXAMPLE

Suppose that the set of social states is given by $X=\{a, b, c, d\}$, and that there are two agents indexed by $i=1,2$. Each agent i has a characteristic θ_i which determines his preferences over the states in X. These preferences are in turn described by a von Neumann-Morgenstern utility function $u_i(\cdot|\theta_i)$ over the set X. Let Θ_i be the set of all possible characteristics for agent i, and suppose that $\Theta_1=\{\theta_1, \theta_1'\}$ and $\Theta_2=\{\theta_2, \theta_2'\}$, where the corresponding utility functions are given by

$$\begin{array}{rcccc} & a & b & c & d \\ u_1(\cdot|\theta_1)= & (2 & 4 & 2 & 4) \\ u_1(\cdot|\theta_1')= & (1 & 0 & 2 & 4) \\ u_2(\cdot|\theta_2)= & (2 & 2 & 4 & 4) \\ u_2(\cdot|\theta_2')= & (1 & 2 & 0 & 4) \end{array}$$

Consider the social choice rule f defined by

$$f=\begin{array}{cc} \begin{array}{cc} \theta_2 & \theta_2' \end{array} & \\ \begin{pmatrix} \{a\} & \{b\} \\ \{c\} & \{d\} \end{pmatrix} & \begin{array}{c} \theta_1 \\ \theta_1' \end{array} \end{array}$$

The social choice rule f can be implemented in dominant strategies by the mechanism

$$g=\begin{array}{cc} \begin{array}{ccc} s_2 & s_2' & s_2'' \end{array} & \\ \begin{pmatrix} a & b & b \\ c & d & c \\ c & b & a \end{pmatrix} & \begin{array}{c} s_1 \\ s_1' \\ s_1'' \end{array} \end{array}.$$

In the mechanism g agent 1 chooses rows as strategies, and agent 2 columns. Agent 1's dominant strategies are s_1 if his characteristic is θ_1 and s_1' if θ_1'. Agent 2's dominant strategies are s_2 if his characteristic is θ_2 and s_2' if θ_2'. It follows then that the mechanism g implements the social choice rule f in dominant strategies.

Consider now the direct mechanism h obtained from g by associating strategies s_1 and s_1' with θ_1 and θ_1', respectively, and strategies s_2 and s_2' with θ_2 and θ_2', respectively, that is

$$h = \begin{array}{cc} \theta_2 & \theta_2' \end{array}$$
$$h = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{array}{c} \theta_1 \\ \theta_1' \end{array}.$$

By the revelation principle (see Theorem 1 below) the mechanism h truthfully implements the social choice rule f in dominant strategies. However, h does not implement f in dominant strategies, because θ_1' and θ_2' are also dominant for agent 1 with characteristic θ_1 and agent 2 with characteristic θ_2, respectively, and yet state d does not belong to the social choice set $f(\theta_1, \theta_2)$.

In what follows we show that the failure of the direct mechanism h to implement the social choice rule f in dominant strategies can be traced back to a problem in the implementation of f by the mechanism g which only becomes apparent when the underlying informational assumptions are explicitly stated.

Consider first the case of complete information, that is, when every agent knows the characteristic of the other agent. It is easy to check that when agents' characteristics are (θ_1, θ_2) the pair (s_1', s_2') is a Nash equilibrium whose outcome is state d. Thus, the mechanism g does not implement the social choice rule f in Nash strategies. Moreover, since

$$u_1(d|\theta_1) = 4 > 2 = u_1(a|\theta_1)$$
$$u_2(d|\theta_2) = 4 > 2 = u_2(a|\theta_2)$$

it follows that the Nash equilibrium (s_1', s_2') Pareto dominates the dominant strategy Nash equilibrium (s_1, s_2).

Next we consider the case of incomplete information, that is, when every agent only knows his own characteristic but has some beliefs about the characteristic of the other agent. For example, suppose that agent 1 (agent 2) assigns probability 1/2 that agent 2's characteristic is θ_2 (agent 1's characteristic is θ_1). Then it is easy to check that apart from the dominant strategy Bayesian equilibrium

$$\sigma_1 = \begin{pmatrix} \overset{\theta_1}{s_1} & \overset{\theta_1'}{s_1'} \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} \overset{\theta_2}{s_2} & \overset{\theta_2'}{s_s'} \end{pmatrix}$$

there is another Bayesian equilibrium, namely

$$\sigma_1' = \begin{pmatrix} \overset{\theta_1}{s_1'} & \overset{\theta_1'}{s_1'} \end{pmatrix} \qquad \sigma_2' = \begin{pmatrix} \overset{\theta_2}{s_2'} & \overset{\theta_2'}{s_2'} \end{pmatrix}$$

in which state d is obtained when agents' characteristics are (θ_1, θ_2). Thus, the mechanism g does not implement the social choice rule f in Bayesian strategies. Moreover, since

$$u_1(d|\theta_1) = 4 > 3 = 1/2 u_1(a|\theta_1) + 1/2 u_1(b|\theta_1)$$
$$u_1(d|\theta_1') = 4 > 3 = 1/2 u_1(c|\theta_1') + 1/2 u_1(d|\theta_1')$$
$$u_2(d|\theta_2) = 4 > 3 = 1/2 u_2(a|\theta_2) + 1/2 u_2(c|\theta_2)$$
$$u_2(d|\theta_2') = 4 > 3 = 1/2 u_2(b|\theta_2') + 1/2 u_2(d|\theta_2')$$

it follows that the Bayesian equilibrium (σ_1', σ_2') Pareto dominates the dominant strategy Bayesian equilibrium (σ_1, σ_2).

3. NOTATION AND DEFINITIONS

Let X denote the set of social states and let $I=\{1,\ldots,n\}$ be the set of agents. Each agent i has a characteristic θ_i which determines a von Neumann–Morgenstern utility function $u_i(\cdot|\theta_i)$ over the set X. Let Θ_i be the set of all possible characteristics for agent i, and let $\Theta=\prod_{i\in I}\Theta_i$.

A *social choice rule* (SCR) is a correspondence $f:\Theta\rightrightarrows X$ which specifies a non-empty choice set $f(\theta)\subset X$ for each vector of characteristics $\theta\in\Theta$.

A *mechanism* is a function $g:S\rightarrow X$ which specifies a social state $g(s)\in X$ for each vector of strategies $s\in S=\prod_{i\in I}S_i$, where S_i denotes agent i's strategy set. A *direct mechanism* is one for which $S_i=\Theta_i$ for all $i\in I$.

Given a mechanism g we say that a strategy $s_i\in S_i$ is dominant for agent i with characteristic θ_i if for all $s'_1\in S_i$ and $s_{-i}\in S_{-i}=\prod_{j\neq i}S_j$ we have

$$u_i[g(s_i, s_{-i})|\theta_i]\geqq u_i[g(s'_1, s_{-i})|\theta_i].$$

Let $D_i(g,\theta_i)\subset S_i$ denote the set of dominant strategies for agent i with characteristic θ_i in the mechanism g, and for each vector of characteristics $\theta\in\Theta$ let $D(g,\theta)=\prod_{i\in I}D_i(g,\theta_i)$.

The mechanism g is said to implement the SCR f in dominant strategies if for all $\theta\in\Theta$:

(i) $D(g,\theta)\neq\emptyset$, and

(ii) $g(D(g,\theta))\subset f(\theta)$

It should be noticed that the concept of implementation in dominant strategies is defined without reference to the information that agents are assumed to have about one another. Once this is introduced we may consider two different informational assumptions, namely, complete and incomplete information. These assumptions will in turn lead to the concepts of implementation in Nash and Bayesian strategies, respectively.

Under complete information it is assumed that each agent i not only knows his own characteristic θ_i, but also the characteristics $\theta_{-i}=(\theta_1,\ldots,\theta_{i-1},\theta_{i+1},\ldots,\theta_n)$ of the other agents. A mechanism g together with a vector of characteristics $\theta\in\Theta$ then defines a game with complete information. A Nash equilibrium for this game is a vector of strategies $s\in S$ such that for all $i\in I$ and $s'_i\in S_i$ we have

$$u_i[g(s_i, s_{-i})|\theta_i]\geqq u_i[g(s'_i, s_{-i})|\theta_i].$$

Let $N(g,\theta)\subset S$ denote the set of Nash equilibria for the game (g,θ).

The mechanism g is said to implement the SCR f in Nash strategies if for all $\theta\in\Theta$:

(i) $N(g,\theta)\neq\emptyset$, and

(ii) $g(N(g,\theta))\subset f(\theta)$.

It is easy to see that the fact that the mechanism g implements the SCR f in Nash strategies does not imply that g implements f in dominant strategies, since $D(g,\theta)$ may be empty for some $\theta\in\Theta$. Also the fact that g implements f in dominant strategies does not imply that g implements f in Nash strategies, since $g(N(g,\theta))$ may not be contained in $f(\theta)$ for some $\theta\in\Theta$.

Under incomplete information it is assumed that each agent i knows his characteristic θ_i, but he does not know the characteristics θ_{-i} of the other agents. Instead he has some beliefs about these characteristics. These beliefs are described by a probability measure $\mu_i(\cdot|\theta_i)$ on the set $\Theta_{-i}=\prod_{j\neq i}\Theta_j$. Thus, in this framework agents' characteristics determine both preferences and beliefs. Finally, it is assumed that the array $(\Theta_i, \mu_i(\cdot|\cdot), \mu_i(\cdot|\cdot))_{i\in I}$ is common knowledge in the sense of Aumann (1976).

A mechanism g together with the array $(\Theta_i, u_i(\cdot|\cdot), \mu_i(\cdot|\cdot))_{i\in I}$ defines a game with incomplete information in the sense of Harsanyi (1967–1968). A strategy rule for agent

i in this game is a function $\sigma_i : \Theta_i \to S_i$ which specifies a strategy choice $\sigma_i(\theta_i) \in S_i$ for each characteristic $\theta_i \in \Theta_i$. Let $\Sigma_i(g)$ be the set of all strategy rules for agent i in the game defined by the mechanism g, and let $\Sigma(g) = \prod_{i \in I} \Sigma_i(g)$.

A Bayesian equilibrium for this game is a vector of strategy rules $\sigma \in \Sigma(g)$ such that for all $i \in I$, $\theta_i \in \Theta_i$, and $s_i \in S_i$ we have

$$\int u_i[g(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i}))|\theta_i] \, d\mu_i\,(\theta_{-i}|\theta_i) \geqq \int u_i[g(s_i, \sigma_{-i}(\theta_{-i}))|\theta_i] \, d\mu_i\,(\theta_{-i}|\theta_i).$$

Let $B(g) \subset \Sigma(g)$ denote the set of Bayesian equilibria for the game defined by the mechanism g, and for each vector of characteristics $\theta \in \Theta$ let $B(g, \theta) = \{s \in S | s = \sigma(\theta)$ for some $\sigma \in B(g)\}$.

The mechanism g is said to implement the SCR f in Bayesian strategies if for all $\theta \in \Theta$:

(i) $B(g, \theta) \neq \varnothing$, and

(ii) $g(B(g, \theta)) \subset f(\theta)$

As in the case of implementation in Nash strategies, it is easy to see that the fact that the mechanism g implements the SCR f in Bayesian strategies does not imply that g implements f in dominant strategies, since $D(g, \theta)$ may be empty for some $\theta \in \Theta$. Also, the fact that g implements f in dominant strategies does not imply that g implements f in Bayesian strategies, since $g(B(g, \theta))$ may not be contained in $f(\theta)$ for some $\theta \in \Theta$.

4. RESULTS

In this section we first introduce an alternative notion of implementation which is available for direct mechanisms, namely, truthful implementation in dominant strategies. We then state a well-known result which guarantees that for any mechanism that implements a social choice rule in dominant strategies there exists a direct mechanism which truthfully implements it in dominant strategies. Finally, we show that if the direct mechanism in question does not implement the social choice rule in dominant strategies, then it must be the case that the original mechanism does not implement it in Nash strategies (under complete information) or in Bayesian strategies (under incomplete information).

The direct mechanism h truthfully implements the SCR f in dominant strategies if for all $\theta \in \Theta$:

(i) $\theta \in D(h, \theta)$, and

(ii) $h(\theta) \in f(\theta)$

It should be noticed that implementation and truthful implementation are quite different concepts. Indeed, the fact that the direct mechanism h truthfully implements the SCR f in dominant strategies does not imply that h implements f in dominant strategies, since $h(D(h, \theta))$ may not be contained in $f(\theta)$ for some $\theta \in \Theta$. Also, the fact that the direct mechanism h implements the SCR f in dominant strategies does not imply that h truthfully implements f in dominant strategies, since there may be some $\theta \in \Theta$ which does not belong to $D(h, \theta)$.

Theorem 1 (revelation principle). *If g is a mechanism that implements the SCR f in dominant strategies, and we define the direct mechanism h by $h(\theta) = g(\sigma(\theta))$, where $\sigma(\theta) \subset D(g, \theta)$ for all $\theta \in \Theta$, then h truthfully implements f in dominant strategies.*

Theorem 1 is proved in Dasgupta, Hammond and Maskin (1979, p. 194).

As noted by Dasgupta, Hammond and Maskin, it is important to realize that the only thing that the revelation principle says is that the direct mechanism h *truthfully* implements the SCR f in dominant strategies. In particular, as the example in Section 2 illustrates, h need not implement f in dominant strategies.

The following results establish that the failure of the direct mechanism h to implement the SCR f in dominant strategies always reveals a problem in the implementation of f by the mechanism g which only becomes apparent when the underlying informational assumptions are explicitly stated.

Consider first the case of complete information, that is, when every agent i knows the characteristics θ_{-i} of the other agents.

Theorem 2. *If g is a mechanism that implements the SCR f in dominant strategies, and we define the direct mechanism h as in Theorem* 1, *then $h(N(h, \theta)) \subset g(N(g, \theta))$ for all $\theta \in \Theta$.*

Proof. For any $\theta \in \Theta$ let $\theta' \in N(h, \theta)$. Then for all $i \in I$ we have

$$u_i[h(\theta'_i, \theta'_{-i})|\theta_i] \geqq u_i[h(\theta_i, \theta'_{-i})|\theta_i],$$

so by the definition of h we get

$$\begin{aligned} u_i[g(\sigma_i(\theta'_i), \sigma_{-i}(\theta'_{-i}))|\theta_i] &\geqq u_i[g(\sigma_i(\theta_i), \sigma_{-i}(\theta'_{-i}))|\theta_i] \\ &\geqq u_i[g(s_i, \sigma_{-i}(\theta'_{-i}))|\theta_i] \end{aligned}$$

for all $s_i \in S_i$ (recall that $\sigma_i(\theta_i) \in D_i(g, \theta_i)$). Hence $\sigma(\theta') \in N(g, \theta)$ and so $h(\theta') = g(\sigma(\theta')) \in g(N(g, \theta))$. ||

Since $D(h, \theta) \subset N(h, \theta)$ for all $\theta \in \Theta$, Theorem 2 implies that for any vector of dominant strategies $\theta' \in D(h, \theta)$ we can always find a Nash equilibrium $s' \in N(g, \theta)$ such that $h(\theta') = g(s')$. Thus if $h(\theta') \notin f(\theta)$ we conclude that the mechanism g does not implement the SCR f in Nash strategies.

Next we consider the case of incomplete information, that is, when every agent i only knows his own characteristic θ_i but has some beliefs $\mu_i(\cdot|\theta_i)$ about the characteristics of the other agents.

Theorem 3. *If g is a mechanism that implements the SCR f in dominant strategies, and we define the direct mechanism h as in Theorem* 1, *then for any set of beliefs we have $h(B(h, \theta)) \subset g(B(g, \theta))$ for all $\theta \in \Theta$.*

Proof. Let $\sigma^* \in B(h)$. Then for all $i \in I$ and $\theta_i \in \Theta_i$ we have

$$\int u_i[h(\sigma_i^*(\theta_i), \sigma_{-i}^*(\theta_{-i}))|\theta_i]\, d\mu_i\, (\theta_{-i}|\theta_i) \geq \int u_i[h(\theta_i, \sigma_{-i}^*(\theta_{-i}))|\theta_i]\, d\mu_i\, (\theta_{-i}|\theta_i),$$

so by the definition of h we get

$$\begin{aligned} &\int u_i[g(\sigma_i(\sigma_i^*(\theta_i)), \sigma_{-i}(\sigma_{-i}^*(\theta_{-i})))|\theta_i]\, d\mu_i\, (\theta_{-i}|\theta_i) \\ &\geqq \int u_i[g(\sigma_i(\theta_i), \sigma_{-i}(\sigma_{-i}^*(\theta_{-i})))|\theta_i]\, d\mu_i\, (\theta_{-i}|\theta_i) \\ &\geqq \int u_i[g(s_i, \sigma_{-i}(\sigma_{-i}^*(\theta_{-i})))|\theta_i]\, d\mu_i\, (\theta_{-i}|\theta_i) \end{aligned}$$

for all $s_i \in S_i$ (recall that $\sigma_i(\theta_i) \in D_i(g, \theta_i)$). Hence $\sigma \circ \sigma^* \in B(g)$ and so $h(\sigma^*(\theta)) = g(\sigma(\sigma^*(\theta))) \in g(B(g, \theta))$ for all $\theta \in \Theta$. $\|$

Since $D(h, \theta) \subset B(h, \theta)$ for all $\theta \in \Theta$, Theorem 3 implies that for any vector of dominant strategies $\theta' \in D(h, \theta)$ we can always find a Bayesian equilibrium $\sigma' \in B(g)$ such that $h(\theta') = g(\sigma'(\theta))$. Thus if $h(\theta') \notin f(\theta)$ we conclude that the mechanism g does not implement the SCR f in Bayesian strategies.

5. SUMMARY AND CONCLUSIONS

This paper can be summarized as follows. First we have stated a well-known result in the literature on implementation, known as the revelation principle, which guarantees that for any mechanism g that implements a social choice rule f in dominant strategies there exists a direct mechanism h which truthfully implements f in dominant strategies. We have then noted that the mechanism h need not implement f in dominant strategies. Finally, we have shown that if this is so then it must be the case that the mechanism g does not implement f in Nash strategies (under complete information) or in Bayesian strategies (under incomplete information). Thus, we conclude that the failure of the direct mechanism h to implement the social choice rule f in dominant strategies can always be explained in terms of a more general problem concerning implementation in dominant strategies, namely, the possible existence of alternative Nash or Bayesian equilibria which yield outcomes outside the social choice set.

This problem is an important one because one cannot simply assume that agents will play a vector of dominant strategies when alternative Nash or Bayesian equilibria exist. For example, if for some reason agents choose to play a vector of non-truthful dominant strategies in the mechanism h which yields an outcome outside the choice set, then the same reason may lead them to play the corresponding Nash or Bayesian equilibrium in the mechanism g. But then one should not claim that the problem of implementing the social choice rule f is solved by the mechanism g, even though g does implement f in dominant strategies.

These observations point out the need to introduce explicitly the information that agents are assumed to have about one another. And in the absence of a theory about how agents select an equilibrium in the game with complete or incomplete information defined by a mechanism, one should require that the set of Nash or Bayesian equilibrium outcomes for each vector of characteristics be contained in the corresponding social choice set. In other words, one should require implementation in Nash or Bayesian strategies. Thus, the only role of dominant strategies would be that of ensuring the existence of direct mechanisms that implement the social choice rules under consideration in Nash or Bayesian strategies.

First version received June 1983; *final version accepted June* 1984 (*Eds.*).

I would like to thank Ken Binmore, Partha Dasgupta, Kevin Roberts and two anonymous referees for helpful comments on an earlier draft of this paper.

REFERENCES

AUMANN, R. J. (1976), "Agreeing to Disagree", *Annals of Statistics*, **4**, 1236-1239.

DASGUPTA, P., HAMMOND, P. and MASKIN, E. (1979), "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility", *Review of Economic Studies*, **46**, 185-216.

HARSANYI, J. C. (1967-68), "Games with Incomplete Information Played by 'Bayesian' Players (Parts I-III)", *Management Science*, **14**, 159-182, 320-334, 486-502.