

# Implementation in Economic Environments with Incomplete Information: The Use of Multi-Stage Games\*

Sandeep Baliga<sup>†</sup>

*Kellogg Graduate School of Management, Department of Managerial Economics and Decision Sciences, 2001 Sheridan Road, Evanston, Illinois 60208-2009*

Received February 22, 1996

This paper shows that, in economic environments with incomplete information, incentive compatibility and a preference reversal condition are sufficient for implementation in sequential equilibrium. *Journal of Economic Literature Classification Numbers: C72, D71, D82.* © 1999 Academic Press

*Key Words:* implementation; incomplete information; multi-stage games; sequential equilibrium; economic environments.

## 1. INTRODUCTION

Implementation theory is concerned with the design of procedures to enable a central authority to realize its goals in an environment where it does not know agents' preferences. One interpretation of the theory is to see the central authority as a social planner who has decided on a method of aggregating agents' preferences into social choices. His problem is to ensure that the outcome which is implemented is optimal with respect to the true preferences, whatever they turn out to be. The procedure or mechanism is then designed so that *all* its equilibria reveal this information and achieve the planner's objectives: full implementation.

This is the notion of implementation we will use.<sup>1</sup> We call a vector of preferences for individuals a preference profile. We are concerned with

\*This is Chapter 1 of my Ph.D. thesis at Harvard University. I would like to thank my advisors Jerry Green, Eric Maskin, and Tomas Sjöström. Sandro Brusco, Boaz Moselle, and Ben Polak were also extremely helpful. An Associate Editor and a referee also helped to significantly improve the paper. Any remaining errors are my responsibility.

<sup>†</sup>E-mail: [baliga@nwu.edu](mailto:baliga@nwu.edu)

<sup>1</sup>An alternative notion, weak implementation, is suggested by the Revelation Principle and considers conditions under which truth-telling is one equilibrium of a revelation game.

implementation of social choice functions (scfs) which identify an outcome (the social optimum) for each feasible preference profile. Following Maskin (1977), much of the literature has focused on the complete information environment where the preference profile is common knowledge among the agents but the planner does not know it. If a purpose of the implementation literature is to study the planner's problem when information is decentralized, considering the impact of the decentralization of agents' information is also important. Therefore, the main objective of this paper is to analyze implementation of social choice functions in incomplete information environments. As we use sequential rationality arguments in the extensive form, this paper can be thought of as an incomplete information counterpart of Moore and Repullo (1988) who consider subgame perfect implementation.<sup>2</sup>

We restrict our attention to the case of private values, independent types, and diffuse information: an agent's utility function depends on his own type but not those of other agents, there is no correlation between the types of different agents, and all types have positive probability. We show that in economic environments where there is a large quantity of a private good in which all agents' utilities are strictly increasing, incentive compatibility and a "preference reversal" condition are sufficient for implementation in sequential equilibrium. Therefore, the Bayesian monotonicity condition necessary for Bayesian implementation can be dispensed with in economic environments.

Much of the research on implementation in incomplete information environments (Abreu and Matsushima, 1992; Jackson, 1991; Mookherjee and Reichelstein, 1990; Palfrey and Srivastava, 1987, 1989a, 1989b; and Postlewaite and Schmeidler, 1986) utilizes normal form games. During the initial circulation of this paper, it was brought to the author's attention that there were two other papers dealing with a similar problem: Bergin and Sen (1997) and Brusco (1995a). Our work is independent. Brusco (1995a) provides necessary and sufficient conditions for implementation in perfect Bayesian equilibrium and Bergin and Sen (1997) provide a sufficient condition for implementation in sequential equilibrium. We discuss their sufficient conditions in the next section. More recently, Brusco (1995b) considers implementation in perfect Bayesian equilibrium in economic environments and in particular in auctions with correlated types. Duggan (1996) considers implementation in extensive form games when agents' quasi-linear preferences are private information.

The next section is devoted to a requisite elucidation of the notation and definitions and Section 3 contains the main result.

<sup>2</sup>Similarly, one can consider what refinements such as perfect equilibrium imply for conditions under which the Revelation Principle holds (Kalai and Samet, 1992).

## 2. NOTATION AND DEFINITIONS

*The Model and Social Optimality*

There is a set of agents,  $N$ , and by a slight abuse of notation, we also represent the cardinality of this set by  $N$ . We assume throughout that  $N \geq 3$ . The set of *outcomes*, the objects of social choice, is  $A$ . An *agent*  $i$ 's type,  $\theta_i \in \Theta_i$ , fully describes his preferences over  $A$  and we assume each  $\Theta_i$  is finite. We assume that preferences satisfy the von Neumann and Morgenstern axioms. It is sometimes convenient to use utility functions,  $u_i(\cdot | \theta_i)$ , defined over  $A$ . It should be clear from the notation that we assume preferences depend only on an agent's own type and not those of other agents. Let  $R^i(\theta_i)$  be the weak preference relation associated with agent  $i$ 's utility function, and  $P^i(\theta_i)$  and  $I^i(\theta_i)$  the corresponding strict and indifference preference orderings.

We assume that at the start of the game, Nature reveals an agent's type to him but that he is given no information about other agents' types. The  $N$ -dimensional vector  $\theta$  is called a *preference profile* or *state*. It is assumed in this paper that there is a common knowledge *prior distribution*,  $p(\cdot)$ , over the set of preference profiles and that all types have positive probability. We assume that types are independent so the probability of a preference profile  $\theta$  is  $p(\theta) = \prod_{i \in N} p_i(\theta_i)$  where  $p_i(\theta_i)$  is the probability that agent  $i$  is of type  $\theta_i$ . An *environment* is a collection  $[N, \{\Theta_i\}_{i=1}^N, A, \{p_i\}_{i=1}^N, \{u_i\}_{i=1}^N]$ . We assume that the structure on an environment is common knowledge among the agents.

We will look at a particular set of environments and assume that there is a perfectly divisible, private good, "money," that is available in unlimited quantities. Transfers of money will only be used "out of equilibrium" to design incentives. We will incorporate them explicitly by redefining the set of outcomes and agents' utility functions appropriately. Denote by  $T_i = \mathfrak{R}$  the set of possible net transfers of money to or from agent  $i$  and let  $\tau_i$  be a generic element of  $T_i$ . Let the set of *extended outcomes* be  $A^* = AxT_1x \cdots xT_N$ . Therefore an outcome is now  $(a, \tau_1, \dots, \tau_N)$  and the utility of agent  $i$  of type  $\theta_i$  is  $u_i(a, \tau_1, \dots, \tau_N | \theta_i)$ . If an outcome  $a \in A$  is implemented with zero net transfers, we write it as  $(a, 0_1, \dots, 0_N)$ .

**DEFINITION.** The environment is *economic* (*Condition E*) if the following conditions hold:

- (1) there is a private good, money, which is available in unlimited quantities;
- (2) there are no externalities between agents as far as these monetary transfers are concerned: for all  $i$  in  $N$ ,  $\theta_i$  in  $\Theta_i$ , and  $a$  in  $A$ ,  $u_i(a, \tau_i, \tau_{-i} | \theta_i) = u_i(a, \tau_i, \tau'_{-i} | \theta_i) = u_i(a, \tau_i | \theta_i)$  for all  $\tau_{-i}$  and  $\tau'_{-i}$ ; and

(3) utility functions are unbounded below and, *ceteris paribus*, each agent strictly prefers more money to less: for all  $i$  in  $N$ ,  $\theta_i$  in  $\Theta_i$ , and  $a$  in  $A$ ,

- (a)  $\lim_{\tau_i \rightarrow -\infty} u_i(a, \tau_i | \theta_i) = -\infty$ ;
- (b) if  $\tau_i > \tau'_i$ , then  $u_i(a, \tau_i | \theta_i) > u_i(a, \tau'_i | \theta_i)$ .<sup>3</sup>

Notice that this definition does not rule out the existence of public goods or externalities.

DEFINITION. A *social choice function*,  $f$ , chooses an outcome in  $A$  for each profile  $\theta \in \Theta$  together with zero net transfers.<sup>4</sup>

This is our notion of social optimality.

DEFINITION. A social choice function  $f$  is *incentive compatible* if and only if

$$Eu_i[f(\theta) | \theta_i] \geq Eu_i[f(\tilde{\theta}_i, \theta_{-i}) | \theta_i] \text{ for all } \theta_i, \text{ for all } \tilde{\theta}_i, \text{ for all } i.$$

This condition says that, given all other agents are telling the truth, each agent has no incentive not to do so, too. It is well known that a scf  $f$  is implementable in incomplete information environments only if it is incentive compatible.

DEFINITION. *Preference reversal (Condition PR)* Given an ordered pair of preference profiles  $\theta$  and  $\phi$  in  $\Theta$  where  $\theta \neq \phi$ , there exists an agent  $j(\theta, \phi)$ , or  $j$  for short, and a pair of outcomes in  $A$ ,  $a(\theta, \phi)$  and  $b(\theta, \phi)$ , or  $a$  and  $b$  for short, such that

$$aP_j(\theta)b \text{ and } bP_j(\phi)a.$$

Similar conditions can be found in literature on implementation when there is complete information (see the third part of the definition of an economic environment in Moore and Repullo, 1988, and Property Q in Palfrey and Srivastava, 1991). The literature on implementation in normal form games when there is incomplete information instead identifies conditions for implementation based on preference reversal over social choice functions (see Jackson, 1991, for example). Brusco (1995a) provides a sufficient condition, sequential monotonicity no veto, for implementation

<sup>3</sup>To simplify the analysis, we make the assumptions that there is large amount of money and that utility functions are unbounded below. These assumptions are stronger than those made in definitions of economic environments offered in Moore and Repullo (1988) and Jackson (1991).

<sup>4</sup>Recall that the transfers we have defined explicitly are used only "out-of-equilibrium" to design incentives. Any other transfers are already implicit in the outcome in  $A$  recommended by the social choice function.

in perfect Bayesian equilibrium in extensive form games in a general setting. His sufficient condition is complex and involves a string of social choice rules and associated beliefs such that a preference reversal occurs at the end of the string. Bergin and Sen (1997) provide a simpler sufficient condition, posterior reversal, for implementation in sequential equilibrium. Posterior reversal includes a preference reversal condition similar to our Condition PR. It also requires, roughly speaking, first, that there exist a "reward" that can be given to a player who reveals non-optimal play in an implementing mechanism and, second, that a player who falsely claims that there is non-optimal play can be punished. If our Condition E holds, these two requirements are automatically satisfied. Therefore, posterior reversal is automatically satisfied in our model.

### *Multi-Stage Games of Observed Actions and Incomplete Information*

The extensive form is in stages indexed by  $t$  and these are finite in number. At each stage, agents send messages simultaneously. Let agent  $i$ 's message at stage  $t$  be  $m_i^t \in M_i^t$ , where at each stage each agent's action set is at most countable. Let  $m^t$  be the profile of messages at state  $t$ . We will be considering multi-stage games of observed actions. Therefore, *history at stage  $t$*   $h^t$ ,  $(m^1, \dots, m^{t-1})$ , is observed by all agents.

A behavior strategy,  $s_i$ , maps the set of possible histories and types into the set of probability distributions over messages:  $s_i(m_i^t | h^t, \theta_i)$  is the probability of  $m_i^t$  given history  $h^t$  and type  $\theta_i$ . Let  $S_i$  be the set of all behavior strategies for agent  $i$ . The *conditional probability distribution* is  $\mu_i(\theta_{-i} | h^t, \theta_i)$ . This is assumed to exist for all  $i$ ,  $\theta_i$ ,  $t$ , and  $h^t$  and is agent  $i$ 's *belief* over the other agents' types at each stage after each history. We also say a vector of beliefs,  $\mu$ , is *Bayes consistent* with a strategy profile,  $s$ , if beliefs are updated from one stage to the next using Bayes' rule whenever it is possible (see Fudenberg and Tirole, 1991b, for a precise definition). Let  $w(s; h^t)$  be the outcome when agents use strategy  $s$  after history  $h^t$ . We can write as  $Eu_i[w(s; h^t) | h^t, \mu_i(\cdot | h^t, \theta_i), \theta_i]$  the expected utility of agent  $i$  of type  $\theta_i$  given that stage  $t$  is reached after history  $h^t$ . By a slight abuse of notation, we will, in fact, write it as

$$Eu_i[s | h^t, \mu_i(\cdot | h^t, \theta_i), \theta_i].$$

### *Sequential Equilibrium*

DEFINITION. A *sequential equilibrium assessment* is a pair  $(s, \mu)$  of strategies and beliefs for all agents such that

- (S) (sequential rationality) for all  $i \in N$ ,  $\theta_i \in \Theta_i$ ,  $s'_i \in S_i$ , and  $h^t$ ,  
 $Eu_i[s | h^t, \mu_i(\cdot | h^t, \theta_i), \theta_i] \geq Eu_i[(s'_i, s_{-i}) | h^t, \mu_i(\cdot | h^t, \theta_i), \theta_i]$ , and

(C) (consistency) there exists a sequence of perfectly mixed strategies  $(s_1^n, \dots, s_N^n)$  converging to  $(s_1, \dots, s_N)$  with *Bayes consistent* beliefs  $(\mu_1^n, \dots, \mu_N^n)$  converging to  $(\mu_1, \dots, \mu_N)$  where a perfectly mixed strategy is one that puts positive probability on all pure strategies.

Note that condition (C) implies the following three restrictions on off equilibrium path beliefs (see Fudenberg and Tirole, 1991a, 1991b, for details).<sup>5</sup>

B(i) Posterior beliefs are independent and all types of agent  $i$  have the same beliefs:

$$\mu_i(\theta_{-i} | h^t, \theta_i) = \prod_{j \neq i} \mu_i(\theta_j | h^t).$$

This requires that unexpected deviations do not lead agent  $i$  to believe that the other agents' types are correlated.

B(ii) Agents use Bayes' rule, whenever possible, to update beliefs from one stage to another. For all  $i, j, h^t$ , and  $m_j^t \in M_j^t$ :

if for  $\tilde{\theta}_j$ ,  $\mu_i(\tilde{\theta}_j | h^t) > 0$  and  $s_i(m_j^t | h^t, \tilde{\theta}_j) > 0$ , then for all  $\theta_j$ ,

$$\mu_i(\theta_j | (h^t, m^t)) = \frac{\mu_i(\theta_j | h^t) s_j(m_j^t | h^t, \theta_j)}{\sum_{\tilde{\theta}_j} \mu_i(\tilde{\theta}_j | h^t) s_j(m_j^t | h^t, \tilde{\theta}_j)}.$$

The expression above shows how agents update probabilities, when they can, from history  $h^t$  to  $h^{t+1} = (h^t, m^t)$ . It requires that if agent  $j$  did not deviate from his prescribed strategy in stage  $t$ , we continue to use the rule when we update information about him. It does not restrict beliefs about an agent who has just deviated but requires that we continue to update using Bayes' rule on newly assigned beliefs.

B(iii) Agents  $i$  and  $j$  have the same beliefs about the type of a third agent  $k$ . For all  $h^t$  and  $i \neq j \neq k \neq i$ :

$$\mu_i(\theta_k | h^t) = \mu_j(\theta_k | h^t).$$

These are necessary conditions for an assessment to be a sequential equilibrium and we will use them to check if certain assessments are sequential equilibria.

Let  $SE(g)$  be the sequential equilibria of the game  $g$  and let  $SE(g, \theta)$  be the set of sequential equilibrium outcomes of  $g$  in state  $\theta$ .

<sup>5</sup>Although their work considers games with a finite number of strategies, the part of their result we use here is true for games with a countable number of strategies.

### Implementation

DEFINITION. A game  $g$  implements  $f$  in sequential equilibrium, if and only if, for all  $\theta \in \Theta$ ,  $SE(g, \theta) = f(\theta)$ . A scf  $f$  is implementable in sequential equilibrium if and only if there exists a game  $g$  that implements  $f$  in sequential equilibrium.

This is the strongest definition of implementation, where we require equivalence between the equilibria of the mechanism and the outcomes chosen by a social choice function.

## 3. A MECHANISM FOR SEQUENTIAL IMPLEMENTATION

We begin by constructing “prizes” which will be used to reward agents for revealing non-optimal play. Let  $R(f) = \{a \in A \mid (a, 0_1, \dots, 0_N) = f(\theta) \text{ for some } \theta \in \Theta\}$ . Given any ordered pair of preference profiles  $\theta$  and  $\phi$  in  $\Theta$  where  $\theta \neq \phi$ , pick one pair of outcomes  $a(\theta, \phi)$  and  $b(\theta, \phi)$  in  $A$  such that Condition PR holds for some agent  $j(\theta, \phi)$ . Let  $R(\text{PR})$  be the set of these picked outcomes for all ordered pairs of preference profiles  $\theta$  and  $\phi$  in  $\Theta$  where  $\theta \neq \phi$ . Notice that as there are a finite number of types for each agent, for all agents  $i$ ,  $u_i(a \mid \theta_i)$  is uniformly bounded for all  $\theta_i \in \Theta_i$  and for all  $a$  in  $R(f) \cup R(\text{PR})$ . Let  $a_0$  be some arbitrary outcome in  $A$ ,  $\tau^*$  be a large, positive transfer and the outcome  $c_i = (a_0, (\tau_i^*, -(\tau^*/(N-1))_{-i})$ . Therefore, for  $\tau^*$  large enough, for all  $i$  in  $N$  and  $\theta_i$  in  $\Theta_i$ ,  $u_i(c_i \mid \theta_i) > u_i(c_j \mid \theta_i)$  for all  $j$  in  $N \setminus \{i\}$  and  $u_i(c_i \mid \theta_i) > u_i(a, 0 \mid \theta_i)$  for all  $a$  in  $R(f) \cup R(\text{PR})$ . Therefore, agent  $i$  prefers the prize outcome  $c_i$  to other agents' prizes, outcomes in the range of the scf or the picked outcomes over which agents experience preference reversal. Let the set of prizes be  $C = \{c_1, c_2, \dots, c_N\}$ .

The implementation mechanism is as follows:

*Stage N:* Nature moves.

*Stage 1:* Agents simultaneously announce their types. Call the ensuing vector of types  $\theta$ .

*Stage 2:* Agents simultaneously either announce a doublet of “Yes” and a non-negative integer  $n^i$ , or a doublet of a preference profile and a non-negative integer,  $n^i$ .

(2.1) If all agents announce “Yes,” implement  $f(\theta)$

(2.2) If  $N - 1$  agents announce “Yes” and one agent  $k^*$  announces a preference profile,  $\phi$ , then

(2.2.1) if  $f(\theta) = f(\phi)$ , implement  $f(\theta)$ ;

(2.2.2) if  $f(\theta) \neq f(\phi)$  and if  $k^* = j(\theta, \phi)$ , implement  $f(\theta)$ ;

(2.2.3) if  $f(\theta) \neq f(\phi)$  and if  $k^* \neq j(\theta, \phi)$ , go to the Substage.

(2.3) If none of the above apply, the agent who announced the highest  $n^i$  is allowed to choose from C. Ties are broken by randomly choosing amongst the agents who are tied so that each of the agents has an equal probability of being picked.

*Substage:* Each agent raises a “flag” or announces a non-negative integer,  $n^i_s$ . The mechanism distinguishes between different cases as follows:

(S.1) If  $N - 1$  or more flags are raised, then agent  $k^*$  is allowed to choose from C.

(S.2) If  $N - 1$  or more agents announce zero, and

(S.2.1) agent  $j(\theta, \phi)$  is not raising a flag, choose  $a(\theta, \phi)$  and set a large fine for agent  $k^*$ ;

(S.2.2) agent  $j(\theta, \phi)$  is raising a flag, then implement outcome  $b(\theta, \phi)$ .

(S.3) If neither (S.1) nor (S.2) apply, the agent who announced the highest integer is allowed to choose from C, where raising a flag counts as announcing  $-1$ . Ties are broken by randomly choosing amongst the agents who are tied so that each of the agents has an equal probability of being picked.

**THEOREM 1.** *Suppose that  $N \geq 3$ . In an environment satisfying Conditions E and PR, any incentive compatible scf is implementable in sequential equilibrium.*

*Proof.* We show that a “truthful” sequential equilibrium implements the scf and that there are no non-optimal sequential equilibria. For simplicity, we will refer to sequential equilibrium (equilibria) as “equilibrium” (“equilibria”). The proof will proceed according to the following steps:

First, we prove a Claim that will be helpful for the proof. Let  $Z(F)$  denote the event that an agent believes that all others will play “zero” (“flag”) in the Substage. The Claim will show that, in an equilibrium, an agent must believe either  $F$  or  $Z$ .

Second, we show that, in any equilibrium, all agents must tell the truth.

Finally, we show that an optimal equilibrium exists by describing equilibrium strategies at all information sets, both on and off the equilibrium path, for all possible preference profiles.

Recall that a sequential equilibrium assessment is a strategy profile and belief pair. The following Claim establishes some restrictions on what beliefs agents can hold if the history reaches the Substage.

*Claim.* In all equilibria, if the history reaches the Substage, either each agent expects all other agents raise a flag,  $F$ , with probability one or to announce zero,  $Z$ , with probability one.



*Proof of Claim.* Suppose not and, in an equilibrium, an agent  $i^*$  expects others either, not to unanimously raise flags with probability one or, not to announce zero with probability one. If agent  $i^*$  is not announcing the highest strictly positive integer there are three possibilities: (1) he is tied for first place when there is some probability that he will get to choose his favorite outcome; (2) some agent  $j$ 's favorite outcome  $c_j$  is implemented; (3) some outcome  $a(\theta, \phi)$  or  $b(\theta, \phi)$  is implemented (together possibly with a fine for some agent). If he announces the highest strictly positive integer, he will definitely get to pick his most preferred outcome from  $C$ . Therefore, in an economic environment, he strictly prefers this strategy whatever his type. Hence, even if he is playing type-dependent actions, whatever beliefs are, other agents will know that he is announcing strictly positive integers. But then, some other agent has an incentive to deviate and announce even higher integers and the Claim is proved.

LEMMA 1. *If  $\theta^*$  is the true preference profile, then it is announced with probability one in any equilibrium.*

*Proof.* Let  $(s, \mu)$  be an equilibrium and let  $\theta^*$  be the true preference profile.

*Step 1:* For all  $i$  in  $N$ ,  $s_i(\theta_i^*)$  must specify that agent  $i$  announces (Yes,.) in Stage 2 after all histories. Any other strategies trigger the integer game which has no equilibrium.

*Step 2:* For all  $i$  in  $N$ ,  $\mu_i$  must specify the belief  $Z$  if the history reaches the Substage.

Suppose not.

*Case 1.* Suppose agent  $i$ 's beliefs  $\mu_i$  specify  $Z$  for some history reaching the Substage but for some agent  $j$ ,  $\mu_j$  specifies  $F$ , the only other possibility in equilibrium by the Claim. By condition B(iii), agent  $i$  and agent  $j$  have the same beliefs about agents in the set  $N \setminus \{i, j\}$  which is non-empty by our assumption that there are at least three agents. But agent  $i$  believes agents in the set  $N \setminus \{i, j\}$  will announce zero and agent  $j$  believes they will raise flags, a contradiction. Therefore, Case 1 cannot occur in equilibrium.

*Case 2.* Suppose  $\mu_i$  specifies  $F$  for all  $i$  in  $N$  for some history that reaches the Substage. This history involves one agent, say  $j$ , announcing  $(\phi, .)$  while all other agents announce (Yes, .). By property B(i), all types of agent  $j$  believe he can pick  $c_j$  if he does this as  $\mu_j$  specifies the belief  $F$  in the Substage following this history. In any equilibrium, as established in step 1, agent  $j$  announces (Yes, .) and some outcome in  $R(f)$  is implemented. But then agent  $j$  has an incentive to deviate so Case 2 cannot arise in equilibrium.

The only remaining possibility is that, as claimed, in an equilibrium  $(s, \mu)$ ,  $\mu_i$  must specify the belief  $Z$  for all  $i$  in  $N$  if the history reaches the Substage.

*Step 3:* Finally, we claim that, given  $\theta^*$  is the true preference profile,  $\theta^*$  is announced with probability one in Stage 1. Suppose not and a preference profile  $\theta \neq \theta^*$  is announced with positive probability in Stage 1. We now consider if it is possible for  $\mu_i$  to specify  $Z$  for all  $i$  in  $N$  after all histories that lead to the Substage (as is required in any equilibrium by step 2). Denote by  $\phi$  a state where  $\theta$  is announced with positive probability. By Condition PR, as  $j(\theta, \phi)$  prefers  $b(\theta, \phi)$  to  $a(\theta, \phi)$ , if history reaches the appropriate Substage, his sequentially rational strategy must specify that he will raise a flag as he believes  $Z$ . All paths from Stage 2 leading to the Substage involve an agent  $k^* \neq j(\theta, \phi)$  announcing  $(\phi, \cdot)$  while all other agents including  $j(\theta, \phi)$  announcing (Yes,  $\cdot$ ). This may involve deviation from a strategy profile where all agents announce (Yes,  $\cdot$ ) in Stage 2 so we are off the equilibrium path in the Substage. However, as  $k^* \neq j(\theta, \phi)$  and as  $j(\theta, \phi)$  does not have to deviate for play to reach the Substage, other agents must be certain that he will not announce zero with probability one by condition B(ii). Hence, it is impossible for  $\mu_i$  to specify the belief  $Z$  for all agents  $i$  in  $N$  at such a Substage. But this contradicts our findings in step 2 and Lemma 1 is proved.

LEMMA 2. *An optimal sequential equilibrium exists.*

*Proof.* The “truth-telling” equilibrium is supported by the following strategies and beliefs:

(i) all agents announce truthfully at all information sets at Stage 1, say “(Yes, 0)” at Stage 2 and “0” if the history ever reaches the Substage; they play optimally given beliefs at all other information sets; and,

(ii) beliefs are Bayes consistent with truth-telling and, in the Substage, which is off the equilibrium path, we specify them to be that the types announced by other agents in Stage 1 are true.

Given the strategies specified, the Substage is off the proposed equilibrium path. We are free to specify beliefs in these subject to our restrictions that they satisfy consistency with the strategies specified. We allow all agents to “tremble” so that they put probability  $1 - \varepsilon$  in all stages on the truth-telling strategies specified above and a total of probability  $\varepsilon$  on all other actions. The beliefs associated with these trembles converge to  $\theta^*$ , the true preference profile. Given these beliefs, it is an equilibrium for all agents to announce zero and believe  $Z$  in the Substage as no agent  $j(\theta, \phi)$  has an incentive to deviate. Therefore, in an economic environment, if the fine is large enough, if the profile  $\theta$  is announced, no agent  $k^* \neq j(\theta, \phi)$

has an incentive to announce a profile  $\phi$  in Stage 2. Finally, as the scf is incentive compatible, all agents weakly prefer to tell the truth to lying in Stage 1 for all states of the world. Q.E.D.

## REFERENCES

- Abreu, D., and Matsushima, H. (1992). "Virtual Implementation in Iteratively Undominated Strategies: Incomplete Information." Mimeo, Princeton University.
- Bergin, J., and Sen, A. (1997). "Extensive Form Implementation in Incomplete Information Environments." Mimeo, Queen's University, Ontario.
- Brusco, S. (1995a). "Perfect Bayesian Implementation," *Econ. Theory* **5**, 429–444.
- Brusco, S. (1995b). "Perfect Bayesian Implementation in Economic Environments." Mimeo, University Carlos III, Madrid.
- Duggan, J. (1996). "An Extensive Form Solution to the Adverse Selection Problem in Principal/Multi-Agent Environments." Mimeo, Queen's University, Canada.
- Fudenberg, D., and Tirole, J. (1991a). "Perfect Bayesian Equilibrium and Sequential Equilibrium," *J. Econ. Theory* **53**, 236–260.
- Fudenberg, D., and Tirole, J. (1991b). *Game Theory*. Cambridge, MA: MIT Press.
- Jackson, M. O. (1991). "Bayesian Implementation," *Econometrica* **59**, 461–477.
- Kalai, E., and Samet, D. (1992). "Types Decomposition and Perfect Implementation," in *Rational Interaction, Essays in Honor of John C. Harsanyi*, Reinhard Selten (Ed.). Berlin: Springer-Verlag.
- Kreps, D., and Wilson, R. (1982). "Sequential Equilibrium," *Econometrica* **50**, 863–894.
- Maskin, E. (1977). "Nash Equilibrium and Welfare Optimality," Department of Economics, MIT.
- Mookherjee, D., and Reichelstein, S. (1990). "Implementation via Augmented Revelation Mechanisms," *Rev. Econ. Stud.* **57**, 453–476.
- Moore, J., and Repullo, R. (1988). "Subgame Perfect Implementation," *Econometrica* **56**, 1191–1220.
- Palfrey, T., and Srivastava, S. (1987). "On Bayesian Implementable Allocations," *Rev. Econ. Stud.* **54**, 193–208.
- Palfrey, T., and Srivastava, S. (1989a). "Mechanism Design with Incomplete Information: A Solution to the Implementation Problem," *J. Polit. Econ.* **97**, 668–691.
- Palfrey, T., and Srivastava, S. (1989b). "Implementation with Incomplete Information in Exchange Economies," *Econometrica* **57**, 115–134.
- Palfrey, T., and Srivastava, S. (1991). "Nash Implementation Using Undominated Strategies," *Econometrica* **59**, 479–501.
- Postlewaite, A., and Schmeidler, D. (1986). "Implementation in Differential Information Economies," *J. Econ. Theory* **39**, 14–33.