Discussion Paper No. 791

FULL BAYESIAN IMPLEMENTATION*

by

Matthew O. Jackson
Kellogg Graduate School of Management
Northwestern University, Evanston IL 60208

First Draft: December 1987
Revised: August 1988

Abstract

In this paper, full implementation is examined for general information structures, including those in which information is incomplete and asymmetric. For a large class of 'economic' environments (including exchange economies), a theorem presents conditions which a collection of state contingent allocations satisfies if and only if there exists a mechanism whose (Bayesian) Nash equilibria exactly coincide with the given allocations. This extends the previous literature both by completely characterizing fully implementable allocations and by doing so for a substantially larger set of environments. A second theorem shows that with the addition of a no veto power condition, the conditions are sufficient for full implementation in a general class of environments in which there are three or more agents. Finally, an example demonstrates a social choice function which is fully implementable when information is incomplete, but which violates monotonicity and is not fully implementable when information is complete.

---

*A portion of this work appeared in a chapter of my dissertation [Jackson (1988)]. I thank the members of my dissertation committee Darrell Duffie, Dilip Mookherjee, and Bob Wilson for their generous help in developing this work. I also thank Faruk Gul, Tom Palfrey, and Andrew Postlewaite for helpful conversations. I wish to acknowledge the support of a National Science Foundation Fellowship, the Graduate School of Business at Stanford University, and a State Farm Doctoral Dissertation Award.

1. Introduction.

A basic issue in the theory of social choice is the question of the existence of a mechanism which decentralizes the process of collective social choice. When agents have differing preferences, Arrow (1963) defined a social choice correspondence to be a map from sets of individual preferences to alternatives, with the idea that the selected alternatives are 'socially preferred' to other alternatives. In this setting, a mechanism is said to implement a given social choice correspondence if its equilibrium outcomes, as a function of agents' preferences, coincide with the social choice correspondence [see Hurwicz (1973)]. With the aim of understanding decentralized collective decision making, we begin by trying to say something about implementation.

Although implementation has been studied most extensively in a full information setting [see Maskin (1985)], it seems to be a more natural issue in an incomplete information setting. In the full information case, it is assumed that every agent has perfect information (while the planner has no information). Perhaps more realistic, is the case in which agents may have incomplete information about their environment. For example, agents might know their own preferences, but only have a conditional probability measure over the preferences of the other agents. Implementation when information is incomplete is the subject of this paper.

In an economy in which agents may have incomplete and asymmetric information about the state of the world, a state–contingent allocation is a map describing the allocation of each agent as a function of the state of the world, and a social choice correspondence is a collection of state contingent allocations. A mechanism is a set of actions and a map from actions into allocations. A mechanism fully implements a social choice correspondence if the set of (Bayesian) Nash equilibria of the mechanism coincides exactly with the social choice correspondence.[1] Characterizing fully implementable social choice correspondences (henceforth abbreviated SCC) amounts to finding a single set of conditions which are satisfied by all fully implementable SCCs and which assure that there exists a mechanism which fully implements a given SCC.

[1] Inherent in our notions of state contingent allocation and social choice correspondence, is the cardinality of agents' preferences. Arrow (1963) did not make cardinal comparisons. Thus, the definition of social choice correspondence which we use differs from Arrow's original definition.

The issue of full implementation when information is incomplete has recently been addressed in several papers. Postlewaite and Schmeidler (1986) and Palfrey and Srivastava (1987) characterize full implementation for exchange economies in which information is 'non–exclusive.' Non–exclusivity requires that the information of any agent be completely redundant, given the information of the other agents in the economy. In fact, under the non–exclusive information assumption, any $N-1$ agents collectively know the exact state of the world. Under this assumption, Postlewaite and Schmeidler (1986) and Palfrey and Srivastava (1987) show that there is a large class of fully implementable allocations. In particular, if information is non–exclusive, it is possible to construct a 'tweed ring' in which every set of $N-1$ agents police the behavior of the remaining agent. Under the non–exclusive information assumption, a Bayesian monotonicity condition [an extension of Maskin's (1977) monotonicity condition] is necessary and sufficient for full implementation, provided $N \geq 3$. Non–exclusivity of information, however, is a restrictive assumption. For example, it rules out any situation in which an agent knows more about the state of the world (or even his own preferences) than the rest of the agents, collectively. It seems that many interesting and realistic situations involve some 'exclusive' information.

In another paper, Palfrey and Srivastava (1985) examine full implementation for exchange economies in which agents may have exclusive information. They show that any fully implementable allocation must satisfy a Bayesian monotonicity condition and a self selection (or Bayesian incentive compatibility) condition. They also show that satisfaction of Bayesian monotonicity and a stronger self selection condition (which they call ϵ–SS) is sufficient for full implementation, when there are three or more agents. However, the stronger self selection condition is not satisfied by some fully implementable allocations.

In this paper, our first theorem presents conditions which are both necessary and sufficient for full implementation in a large class of environments (which includes exchange economies) with general information structures. Specifically, versions of a closure condition, Bayesian monotonicity, and self selection are shown to be both necessary and sufficient for full implementation. This is stronger than the results in the previous literature since it provides a complete characterization of full implementation for general information structures, and since it does so for a larger class of environments (including those with externalities).

A second theorem shows that self selection, closure, and a condition which combines monotonicity and weak no veto power are sufficient for full implementation when we have dropped the assumption that the environment is economic. This covers any environment with any information structure such that there are three or more agents, the set of feasible allocations includes a common worst element and each agent's information set includes knowledge of what his or her most preferred allocation is.

Based on this theory, we are able to present an example showing some of the differences between implementation when information is incomplete and when information is complete. More precisely, we demonstrate a social choice correspondence which is fully implementable when information is incomplete, but which is *not* fully implementable when information is complete. The social choice function violates the monotonicity condition which is necessary for full implementation with complete information. The intuition behind the example is that when information is incomplete, there exists an agent who is able to *credibly* 'fink' on the other agents if they deviate, while when information is complete there is no such *credible* agent. In this case, the additional knowledge changes an agent's incentives and hence credibility.

The paper proceeds as follows. First, we define the notions of environment, social choice correspondence, and full implementation. The first theorem states that, for economic environments, a SCC satisfies closure, Bayesian monotonicity, and self selection, if and only if it is fully implementable. Two subsequent remarks help us to identify the role of the conditions in assuring that a given SCC is fully implementable. If a given SCC satisfies self selection, then there exists a mechanism such that the given SCC is a subset of the equilibria of the mechanism. Similarly, if a given SCC satisfies Bayesian monotonicity and closure, then there exists a mechanism such that the equilibria of the mechanism are a subset of the given SCC. The second theorem generalizes to non–economic environments, presenting sufficient conditions for full implementation in a general class of environments with three or more agents. Next, an example is presented of an allocation which is fully implementable when information is incomplete, but not fully implementable when information is complete. Finally, concluding remarks are made about some of the assumptions and conditions used in the theory. An appendix contains proofs omitted from the text.

2. Definitions.

In general, for a vector (or vector of functions) of dimension n, $v = (v^1, \ldots, v^n)$, we will let $v^{-i} = (v^1, \ldots, v^{i-1}, v^{i+1}, \ldots, v^n)$ and $v/\tilde{v}^i = (v^1, \ldots, v^{i-1}, \tilde{v}^i, v^{i+1}, \ldots, v^n)$.

ENVIRONMENTS

There are a finite number, N, of agents. We use i, j, and k to denote individual agents.

The set S^i describes the finite number of possible information sets of agent i. A *state* is a vector $s = (s^1, \ldots, s^N)$ and the set of states is $S = S^1 \times \cdots \times S^N$. A state provides a full description of agents' preferences and information.[2,3] We will use s and t to denote generic elements of S.

The set of feasible allocations is assumed to be fixed across states and is denoted $\mathcal{A}$.

The state dependent utility of agent i is given by $U^i : \mathcal{A} \times S \to \mathbb{R}_+$.

The following two assumptions are made about an environment.

(A1) $\exists 0 \in \mathcal{A}$ such that $U^i(0, s) = 0$ for all i, s, and for each $a \in \mathcal{A}$ $(a \neq 0)$ and s there exists an agent j such that $U^j(a, s) > 0$, and

(A2) For all i and $s^i \in S^i$, there exists $w^i(s^i) \in \mathcal{A}$ such that $U^i[w^i(s^i), t] \geq U^i[a, t]$ for all $a \in \mathcal{A}$ and $t \in S$ such that $t^i = s^i$.

Assumption (A1) requires that there is a unique alternative which is always the worst for all agents. It is not required that 0 ever be chosen, simply that the mechanism designer have the ability to punish all agents simultaneously. Assumption (A2) requires that for each agent i and possible observation of the agent s^i, there exists an allocation which the agent (weakly) prefers to any other allocation, for any state consistent with the observation. In other words, an agent always knows his or her most preferred allocation. [the necessity of these assumptions is discussed at the conclusion of this paper.]

---

[2] An $s^i \in S^i$ may also be thought of as an agent's type. It must be noted, however, that this is at an interim stage. We will allow agents' preferences to depend on s and hence ex–post, s^i may not fully describe agent i.

[3] Our definition of a state is similar to Moore and Repullo (1986) and Mookherjee and Reichelstein (1987), but different from that used by Postlewaite and Schmeidler (1986) and Palfrey and Srivastava (1985) and (1987). In their setups, a state is a primitive notion and agents information sets are functions of the state. Here, we start with agents information sets and then define a state to be a collection of these information sets. In our study we will treat situations in which agents announce observations which may be incongruent. In our setup this is much easier. Our setup is completely equivalent to theirs and our results can be stated in their framework, but with some loss of brevity.

Each agent i has a prior (probability measure) q^i on the possible states. It is assumed that if $q^i(s) > 0$ for some i and s, then $q^j(s) > 0$ for all $j \neq i$. That is, agents agree on the set of states which occur with positive probability, which is denoted T, $T = \{s \mid q^i(s) > 0,\ \forall i\}$. The measures q^i define partitions Π^i over the set T which are referenced by π^i, $\pi^i(s^i) = \{t : t^i = s^i, q(t) > 0\}$. Thus for a given information set s^i, $\pi^i(s^i)$ defines the set of states which i believes may be the true state. It is assumed without loss of generality that $\pi^i(s^i) \neq \emptyset$ for each i and s^i.

An *environment* is a collection $[S, N, \mathcal{A}, U, q]$. It is assumed that the structure of an environment is common knowledge among the agents.

SOCIAL CHOICE CORRESPONDENCES (SCCs)

Define the set of *feasible state contingent allocations*

$$X = \{x \mid x : S \to \mathcal{A}, \}.$$

A *social choice correspondence* (henceforth abbreviated SCC) is a subset F of X. For the issue of implementation, a SCC may be thought of as a collection of maps which list desirable allocations as a function of the state. [Of course, implicit in the word desirable is some notion of societal value, which we take as given.]

This is a bit different from the definition of a social choice correspondence used when information is complete (a correspondence from states to allocations). The definition used here is consistent with the literature on Bayesian implementation. Given our incomplete information structure and our interest in full implementation, it is natural to associate a state–contingent allocation with each set of equilibrium strategies. Given closure (a necessary condition for implementation), the two definitions are equivalent when information is complete.

Given $x, y \in X$ and $C \subset S$, we define the *weak preference relation* R^i for agent i by

$$x\ R^i(C)\ y \Longleftrightarrow$$

$$\sum_{s \in C} q^i(s) U^i[x(s), s] \geq \sum_{s \in C} q^i(s) U^i[y(s), s].$$

We say that i *strictly prefers* x to y on C, $x\ P^i(C)\ y$, if it is not true that $y\ R^i(C)\ x$.

IMPLEMENTATION

A *mechanism* is an action space $M = M^1 \times \cdots \times M^N$ and a payoff function $g : M \to \mathcal{A}$.

A *strategy* is a mapping $\sigma^i : S^i \to M^i$. Let σ denote the vector of strategies $[\sigma^1, \ldots, \sigma^N]$. We will write $\sigma(s)$ to represent $\left(\sigma^1(s^1), \ldots, \sigma^I(s^N)\right)$. Let $g(\sigma)$ denote the allocation which results when σ is played. A strategy σ^i is a *best response* to σ^{-i} if, for all s and $\widetilde{\sigma}^i$,

$$g(\sigma) \; R^i[\pi^i(s^i)] \; g(\sigma/\widetilde{\sigma}^i).$$

A vector of strategies σ is a Bayesian Nash *equilibrium* if for all i, σ^i is a best response to σ^{-i}.

A mechanism (M, g) *fully implements* a SCC F if

(i) for any $x \in F$ there exists an equilibrium σ with $g[\sigma(s)] = x(s)$ for all $s \in T$, and

(ii) for any equilibrium σ there exists $x \in F$ with $g[\sigma(s)] = x(s)$ for all $s \in T$.

A SCC F is *fully implementable* if there exists a mechanism (M, g) which fully implements F.

Our concern for implementation is on the set T (recall that $T = \{s : \, q(s) > 0\}$). That is, we are concerned with states which occur with positive probability. Notice, however, that state–contingent allocations are defined on all of S. Since our definition of implementation applies only to $s \in T$, any two state–contingent allocations which agree on T are equivalent. Hence, if we begin with a state–contingent allocation x defined on T, any extension of x to all of S is suitable for our study. In general, for a given x [defined on T or on all of S], an extension of x is any $\widehat{x} \in X$ such that $\widehat{x}(s) = x(s)$ for $s \in T$. Given x defined on T or S, let x^0 denote the extension of x such that

$$x^0(s) = x(s) \quad \forall s \in T,$$

$$x^0(s) = 0 \qquad \forall s \notin T.$$

Full implementation requires that the equilibria of the mechanism exactly coincide with the given SCC. The notion of full implementation is stronger than other notions of implementation (for example, truthful implementation). An advantage of the stronger requirement is that it does not allow for the existence of any undesirable equilibria as a by–product of the design. [See Postlewaite and Schmeidler (1986) and Palfrey and Srivastava (1985) for discussion of this point.] Furthermore, as a bonus, this definition of implementation will permit a complete characterization of the set of Bayesian Nash equilibria of games with three or more players in an economic environment.

SELF SELECTION

For an allocation $x \in X$ and an information set $t^i \in S^i$ we define x_{t^i} by

$$x_{t^i}(s) = x(s/t^i), \; s \in S.$$

A SCC, F, satisfies *self selection* (SS) if for all i, $s^i \in S^i$, $x \in F$, and $t^i \in S^i$,

$$x \; R^i[\pi^i(s^i)] \; [x^0]_{t^i}.$$

Self selection is also often referred to as (Bayesian) incentive compatibility. [See Dasgupta, Hammond and Maskin (1979) and Harris and Townsend (1981).] It is well known that an implementable SCC must satisfy (SS). An outcome which we wish to be an equilibrium outcome of some game for a given state, should be weakly preferred to the worst any agent can get from behaving as if another state has occurred. This is clear in the proof of the following lemma.

LEMMA 1. *A fully implementable SCC satisfies (SS).*

PROOF: Let F be a fully implementable SCC and suppose that F does not satisfy (SS). Thus there exist i, $\bar{s}^i$, $x \in F$, and $t^i \in S^i$, such that

$$[x^0]_{t^i} \; P^i[\pi^i(\bar{s}^i)] \; x.$$

Since F is fully implementable there exists a mechanism (g, M) and an equilibrium strategy σ such that $g[\sigma(s)] = x(s)$ for all $s \in T$. Let $\tilde{\sigma}^i(s^i) = \sigma^i(t^i)$ for all $s^i \in S^i$. Since $[x^0]_{t^i}(s) = g[\sigma/\tilde{\sigma}^i(s)]$ for all s such that $s/t^i \in T$ and $[x^0]_{t^i}(s) = 0$ otherwise, it follows that

$$g(\sigma/\tilde{\sigma}^i) \; R^i[\pi^i(\bar{s}^i)] \; [x^0]_{t^i}.$$

Hence,

$$g(\sigma/\tilde{\sigma}^i) \; P^i[\pi^i(\bar{s}^i)] \; g(\sigma).$$

This contradicts the fact that σ is an equilibrium. Hence our supposition was wrong. ∎

BAYESIAN MONOTONICITY

A *deception* for i is a mapping $\alpha^i : S^i \to S^i$. Let $\alpha = (\alpha^1, \dots, \alpha^N)$ and $\alpha(s) = [\alpha^1(s^1), \dots, \alpha^N(s^N)]$. For example, in a direct revelation game α^i would indicate i's announced information set as a function of the true information set.

Consider any deception α and any allocation $x \in F$. A SCC F satisfies *Bayesian monotonicity* if, whenever there is no $z \in F$ such that $z(s) = x^0[\alpha(s)]$ for all $s \in T$, there exist i, $s \in T$, and $y \in X$ such that

$$y \circ \alpha \; P^i[\pi^i(s^i)] \; x^0 \circ \alpha, \text{ while}$$
$$x \; R^i[\pi^i(t^i)] \; y_{\alpha^i(s^i)} \quad \forall t^i \; s.t. \; \alpha(s)/t^i \in T.$$

F satisfies Bayesian monotonicity if it satisfies Bayesian monotonicity if for every α, it satisfies Bayesian monotonicity for α.

Bayesian monotonicity is an extension of Maskin's monotonicity. The condition is discussed in Postlewaite and Schmeidler (1986) and in further detail in Palfrey and Srivastava (1985). The following interpretation of Bayesian monotonicity is a bit different from theirs. The interpretation presented here is based on a closely related condition called selective elimination, which appears in Mookherjee and Reichelstein (1987). In essence, the Bayesian monotonicity condition allows us to 'selectively eliminate' undesirable equilibria. If we consider a mechanism (g, M) which might implement F, and we start with some $x \in F$, then there should be an equilibrium σ such that $g[\sigma(s)] = x(s)$ for all $s \in T$. Without loss of generality, we can assume that $g[\sigma(s)] = x(s)$ for all $s \in S$.[4] Suppose that agents use a deception α, so that the strategies $\sigma \circ \alpha$ are played. The outcome is then $x \circ \alpha$. If there is no $z \in F$ such that for all $s \in T$, $x[\alpha(s)] = z(s)$, then we need to be able to rule out $\sigma \circ \alpha$ as an equilibrium.[5] The Bayesian monotonicity condition is a statement that, in this case, we can rule out $\sigma \circ \alpha$ as an equilibrium. The existence of y with the stated properties allows agent i to signal to the mechanism that α is being played and be rewarded according to y and which makes him better off:

$$y \circ \alpha \; P^i[\pi^i(s^i)] \; x \circ \alpha.$$

[4] This follows from our discussion following the definition of implementation. Any extension of x to $s \notin T$ is equivalent for the study of implementation.

[5] The statement of the condition is in terms of $x^0 \circ \alpha$, since we will use the x^0 extension in our constructive proof. The condition also holds for $x \circ \alpha$ when there exists a σ such that $g(\sigma) = x$ for all $s \in S$.

The other part of the Bayesian monotonicity condition states that

$$x\ R^i[\pi^i(t^i)]\ y_{\alpha^i(s^i)}\ \ \forall t : \alpha(s)/t^i \in T,$$

which assures that if agent i were to falsely accuse the other agents of deceiving [i.e. $\alpha(s)/t^i$ was the true state], then i would not be better off. The condition is further illustrated in the proof of Lemma 2 and in Example 1.

The version of Bayesian monotonicity presented here is stronger than the one used by Palfrey and Srivastava (1985), since it is required to hold for any deception while the condition of Palfrey and Srivastava must only hold for compatible deceptions. [Compatible refers to deceptions such that $\alpha(s) \in T$ whenever $s \in T$.] As the following lemma shows, however, the stronger condition is a necessary condition for implementation.

LEMMA 2. *A fully implementable SCC satisfies (BM).*

PROOF: Suppose that F is implemented by the mechanism (g, M). Take any $x \in F$ and some equilibrium σ such that $g[\sigma(s)] = x(s)$ for all $s \in T$. Suppose that there does not exist a $z \in F$ such that $z(s) = x^0[\alpha(s)]$, for all $s \in T$. Let $\widehat{x}$ be the extension of x such that $g[\sigma(s)] = \widehat{x}(s)$ for all $s \in S$. We show that in each of the following two (exhaustive) cases there exists a $y \in X$ with the desired properties.

CASE 1: There is no $z \in F$ such that $\widehat{x}[\alpha(s)] = z(s)$ for all $s \in T$.

In this case, it must be that $\sigma \circ \alpha$ is not an equilibrium at some $\bar{s} \in T$. Therefore, there exist i and $\bar{m}^i \in M^i$ such that

$$g(\sigma/\widetilde{\sigma}^i \circ \alpha)\ P^i[\pi^i(\bar{s}^i)]\ g(\sigma \circ \alpha),$$

where $\widetilde{\sigma}^i(s^i) = \bar{m}^i$ for all $s^i \in S^i$. Let $y = g(\sigma/\widetilde{\sigma}^i)$. From above,

$$y \circ \alpha\ P^i[\pi^i(\bar{s}^i)]\ \widehat{x} \circ \alpha.$$

Therefore,

$$y \circ \alpha\ P^i[\pi^i(\bar{s}^i)]\ x^0 \circ \alpha.$$

Since $\widetilde{\sigma}^i$ is constant, we know that $y_{\alpha^i(\bar{s}^i)} = y = g(\sigma/\widetilde{\sigma}^i)$. Thus, since σ is an equilibrium we know that

$$x\ R^i[\pi^i(t^i)]\ y_{\alpha^i(\bar{s}^i)}.$$

The allocation y has the properties required in the statement of (BM).

CASE 2: $\widehat{x}[\alpha(s)] = z(s)$ for all $s \in T$, for some $z \in F$.

In this case, since by assumption $x^0[\alpha(\bar{s})] \neq z(\bar{s})$ for some $\bar{s} \in T$, from our definitions of x^0 and $\widehat{x}$ it follows that $x^0[\alpha(\bar{s})] = 0 \neq \widehat{x}[\alpha(\bar{s})]$. Furthermore, $x^0[\alpha(s)] = 0$ whenever $x^0[\alpha(s)] \neq \widehat{x}[\alpha(s)]$ for some $s \in S$. Hence,

$$\widehat{x} \circ \alpha \; P^i[\pi^i(\bar{s}^i)] \; x^0 \circ \alpha,$$

for some i. Since σ is an equilibrium, it follows that

$$\widehat{x} \; R^i[\pi^i(t^i)] \; \widehat{x}_{\alpha^i(\bar{s}^i)},$$

for all t^i. Since x and $\widehat{x}$ agree on T, it follows that

$$x \; R^i[\pi^i(t^i)] \; \widehat{x}_{\alpha^i(\bar{s}^i)},$$

for all t^i. Thus the allocation $\widehat{x}$ satisfies the properties required of y in (BM). ∎

CLOSURE

Recall that $T = \{s : \; q(s) > 0\}$, $\pi^i(s^i) = \{t \mid t \in T, \; t^i = s^i\}$, and that the sets π^i form a partition Π^i over T. Let Π denote the common knowledge concatenation defined by $\Pi^1, \ldots, \Pi^N$, that is, Π is the finest partition which is coarser than each Π^i.

Pick any $P_1 \subset \Pi$, and $P_2 \subset \Pi$ such that $P_1 \cup P_2 = T$, and let $p_1 = \{s \mid \exists p \in P_1, s \in p\}$ and $p_2 = \{s \mid \exists p \in P_2, s \in p\}$. The sets p_1 and p_2 are such that every agent always knows in which set the state of the world lies.

A SCC satisfies *Closure* (C) if for any $x \in F$ and $y \in F$, there exists a $z \in F$ such that $z(s) = x(s) \;\; \forall s \in p_1$, and $z(s) = y(s) \;\; \forall s \in p_2$.

Closure is a fairly basic requirement for fully implementable correspondences. Consider a given SCC, which we wish to implement. Suppose there exists a game which implements the SCC, and suppose, for instance, that the common knowledge concatenation Π has two elements. Consider any two (possibly the same) equilibrium strategies to the game. We know that a third strategy, in which agents act according to the first strategy on the first element of Π and the second strategy on the second element of Π, must also be an equilibrium. Hence, if the SCC is fully implementable, it must contain the outcome of the third strategy. Therefore, it must satisfy closure.

LEMMA 3. *A fully implementable SCC satisfies (C).*

PROOF: Pick any x and y in F and any p_1 and p_2 as described in the definition of (C). Let σ_1 and σ_2 be the equilibrium strategies yielding x and y, respectively. Define σ_3 by $\sigma_3^i(t^i) = \sigma_1^i(t^i)$ for all $t \in p_1$, and $\sigma_3^i(t^i) = \sigma_2^i(t^i)$ for all $t \in p_2$, for all i. It follows that σ_3 is an equilibrium. Since F is fully implementable, there exists $z \in F$, such that $z(s) = g[\sigma_3(s)]$ for all $s \in T$. It follows that z satisfies the requirements of closure. ▌

3. Full Implementation.

Theorem 1 will show that if we restrict our attention to economies which are 'economic' in nature, then the conditions (SS), (BM), and (C), are both necessary and sufficient for implementation.

Our meaning of 'economic' is summarized by the following condition.

(E) For any $y \in X$ and $s \in S$ there exist i and j $(i \neq j)$ such that $U^i[w^i(s^i), s] > U^i[y(s), s]$ and $U^j[w^j(s^j), s] > U^j[y(s), s]$, where $w^i(s^i)$ is as defined in (A2).

Condition (E) requires that for any given allocation and state, there are always two agents who prefer some other allocation. Condition (E) is 'economic' in nature since it implies that there are competing interests among the agents. For instance, notice that assumptions (A1) and (A2), and condition (E) are satisfied by an exchange economy where $N \geq 3$, the set of allocations are a subset of a Euclidean space having maximal and minimal elements, and agents' utility functions are strictly increasing in all states.[6] Condition (E) does not, however, limit us to Arrow–Debreu type settings. It applies to a large class of environments, including those in which there are public goods and/or externalities. Basically, (E) captures all the environments for which we do not need some sort of no veto condition.

Notice that together (A2) and (E) imply that $N \geq 3$. To see this, let $y = w^1$ and then pick any $s \in T$ and apply (E). There must exist two agents who prefer their $w^i(s^i)$ to $w^1(s^1)$.

We are now ready to state our first result of the paper.

THEOREM 1. *In an environment which satisfies (E), a SCC is fully implementable if and only if it satisfies (SS), (BM), and (C).*

---

[6] Thus, the exchange economy settings in Postlewaite and Schmeidler (1986) and Palfrey and Srivastava (1985) and (1987) are special cases of the environments considered here.

The proof of Theorem 1 is constructive. That is, for a given F which satisfies (SS), (BM), and (C), we provide an explicit mechanism which implements F. The proof is contained in the Appendix.

Before we examine general environments, let us identify the roles of the conditions in Theorem 1. The following remarks show us the relationship between the conditions and other notions of implementation. Recall that (g, M) fully implements F if and only if

(i) for any $x \in F$, there exists an equilibrium σ to (g, M) such that $g[\sigma(s)] = x(s)$ for all $s \in T$, and

(ii) for any equilibrium σ of (g, M), there exists $x \in F$ such that $g[\sigma(s)] = x(s)$ for all $s \in T$.

We say that (g, M) *truthfully implements* a given SCC F if (i) holds. Likewise, we say that (g, M) *weakly implements* a given SCC F if (ii) holds.

REMARK 1. *If a given SCC F satisfies (SS), then it is truthfully implementable.*

Remark 1 is a simple generalization of a similar theorem (applying to allocations) in Harris and Townsend (1981). Remark 1 requires only that $N \geq 2$. From Theorem 1 and Remark 1, we know that (BM) and (C) are important in assuring that all the equilibria of an implementing mechanism lie within the given SCC. In fact:

REMARK 2. *In an environment which satisfies (E), if a given SCC F satisfies (BM) and (C), then it is weakly implementable.*

The proofs of the two remarks are at the end of the Appendix.

4. Implementation in General Environments.

We now drop the assumption that (E) holds.

We begin by defining a condition which combines the notions of Bayesian monotonicity and no veto power [see Maskin (1985)].

MONOTONICITY–NO VETO

We say that a state contingent allocation $z \in X$ satisfies the weak no veto hypothesis (NVH) at $s \in T$, if there exists i such that for all $j \neq i$

$$U^j(z(s), s) = U^j(w^j(s^j), s).$$

The no veto hypothesis identifies situations in which at least $N-1$ agents find an allocation most desirable. The statement given above is slightly stronger (but easier to read) than what we will need in Theorem 2. The Theorem would still hold if (NVH) was modified so that each $j \neq i$ finds the allocation to be most desirable for all $t \in pi^i(s^i)$.[7]

Consider any deception α, $z \in X$, and $C \subset T$, such that z satisfies (NVH) for all $s \in T$, $s \notin C$, and for each $s \in C$ there exists $x_s \in F$ such that $z(t) = x_s^0[\alpha(t)]\ \forall t \in C \cap [\cup_i \pi^i(s^i)]$. A social choice correspondence satisfies *monotonicity–no veto* (MNV) if whenever there is no $\hat{z} \in F$ such that $\hat{z}(s) = z(s)$ for all $s \in T$, there exist i, $y \in X$ and $s \in C$ such that

$$y \circ \alpha\ P^i[\pi^i(s) \cap C]\ z, \text{ while}$$

$$x_s\ R^i[\pi^i(t^i)]\ y_{\alpha^i(s^i)}\quad \forall t^i \text{ s.t. } \alpha(s)/t^i \in T.$$

The above condition (MNV) reduces to (BM) if $C = T$ and it reduces to a pure no veto condition when $C = \emptyset$. When information is complete, the (MNV) condition is equivalent to the monotonicity and no veto conditions used by Maskin (1985). Although in its 'combined' form, the condition is a bit lengthy, it is weaker than what would be needed if we were to split the condition into a monotonicity part and a no veto part. This follows since the portions of T on which an allocation satisfies (NVH) may not fall along the common knowledge concatenation of T. That is that it may satisfy (NVH) on only a portion of some $\pi^i(s^i)$ for some i and s^i. Thus, in order to be able to piece together separate monotonicity and no veto conditions, we would have to either require a stronger closure condition or weaken our requirement of implementation.[8]

We are now ready to state our second result.

THEOREM 2. *If a social choice correspondence satisfies (C), (SS), and (MNV), then it is fully implementable.*

The proof of Theorem 2 appears in the Appendix.

The following simple example illustrates the theory developed in Theorems 1 and 2 and provides some insight into the differences between implementation when information

[7] The stricter requirement on z [a stricter (NVH)] will produce a weaker condition (MNV), since (MNV) will have to be satisfied less often.

[8] The reason this difficulty is not present when information is complete, is that then each state is an element of the common knowledge concatenation of T, that is $\pi^i(s^i)$ is always a singleton.

is complete and when it is not. Palfrey and Srivastava (1985) provide examples of correspondences which are implementable under complete information but not implementable under incomplete information. Example 1 does the opposite. It provides a SCC which is implementable when information is incomplete, but not when information is complete.

The example also shows that the class of environments which satisfy (E) [and thus are treated by Theorem 1] is reasonably large and includes interesting environments which are not exchange economies. The naming of the allocations in the example is deliberately vague, so that the environment described may have many interpretations. For example it might be an economy [it satisfies (E)] or it might be a description of preferences over political candidates.

EXAMPLE 1

Consider an environment in which three agents are interacting to select allocations from the set $\mathcal{A} = \{0, a, b, c\}$. The set of possible states is given by $S = \{1, 2\}$. The information structure is given by $S^1 = \{1\}$, $S^2 = \{1\}$, and $S^3 = \{1, 2\}$. The state $s = 1$ refers to the combination $(s^1, s^2, s^3) = (1, 1, 1)$ and the state $s = 2$ refers to $(s^1, s^2, s^3) = (1, 1, 2)$. Thus, agent 3 knows the state and the other two agents do not. The priors of the three agents are identical and given by $q(1) = 1/3$ and $q(2) = 2/3$. The preferences of the agents are represented below.

TABLE 1

U^1	U^2	$U^3(1)$	$U^3(2)$
a	b	c	c
b	a	a	a,b
c	c	b	-
0	0	0	0

A higher position in a given column of the table, indicates a strict preference, while the a, b entry indicates indifference. For example, $U^3(c, 2) > U^3(b, 2) = U^3(a, 2) > U^3(0, 2) = 0$.

Consider the social choice function $x(1) = a$ and $x(2) = b$.[9] We first verify that $F = x$ is fully implementable. Since the environment satisfies (E), the condition (MNV) is equivalent

[9] For appropriate values assigned to the functions U^i, the social choice function described above is utilitarian.

to (BM) and both Theorems 1 and 2 apply. It is clear that F satisfies (C) [F is a function] and (SS) [only agent 3 has more than one information set and (SS) is easily checked for 3]. We now verify that $F = x$ satisfies (BM). There are only three possible deceptions (called α, α', and α'') in this situation. They are described by $\alpha(1) = 2$, $\alpha(2) = 2$; $\alpha'(1) = 1$, $\alpha'(2) = 1$; and $\alpha''(1) = 2$, $\alpha''(2) = 1$. Since $x \circ \alpha \notin F$ for each of the three deceptions, we must be able to find, for each α, an i, $s \in T$, and $y \in X$ such that $y \circ \alpha P^i[\pi^i(s)] x \circ \alpha$, while $x R^i[\pi^i(t^i)] y_{\alpha^i(s^i)}$ for all $t^i : \alpha(s)/t^i \in T$. If i is 1 or 2, since S^i is a singleton, the requirement simplifies to finding a y such that

$$(1/3)U^i(y[\alpha(1)]) + (2/3)U^i(y[\alpha(2)]) > (1/3)U^i(x[\alpha(1)]) + (2/3)U^i(x[\alpha(2)]), \text{ and}$$

$$(1/3)U^i(a) + (2/3)U^i(b) \geq (1/3)U^i(y[1]) + (2/3)U^i(y[2]).$$

Let us first consider α'. The allocation $y(1) = b$ and $y(2) = a$ satisfy the requirement when $i = 2$. It is easily checked that

$$U^2(b) > U^2(a), \text{ and}$$

$$(1/3)U^2(a) + (2/3)U^2(b) \geq (1/3)U^2(b) + (2/3)U^2(a).$$

Similarly, it is easily verified that for α, $i = 1$ and y, where $y(1) = 0$ and $y(2) = a$, satisfy the requirement. For α'', $i = 3$, $s = 1$, and $y(1) = y(2) = a$ satisfy the requirement.

We have thus established that we can fully implement the SCC which consists of the state–contingent allocation $x(1) = a$ and $x(2) = b$. Let us now examine the situation in which information is complete, that is, where agents 1 and 2 also know the state. In this case, $S^1 = \{1,2\}$, $S^2 = \{1,2\}$, and $S^3 = \{1,2\}$. The state $s = 1$ refers to the combination $(s^1, s^2, s^3) = (1,1,1)$ and the state $s = 2$ refers to $(s^1, s^2, s^3) = (2,2,2)$. Here, $q(1) = 1/3$, $q(2) = 2/3$ and $q(s) = 0$ for any other s. We verify that (BM) is violated. It is violated at α', where $\alpha'(s) = (1,1,1)$, for all s. For one of agents 1 and 2 to satisfy the requirement for Bayesian monotonicity it would be necessary that

$$U^i(a) \geq U^i(y(1)), \text{ and } U^i(y(1)) > U^i(a),$$

which can never be satisfied. Thus the requirement, if met, would have to be satisfied by agent 3. Therefore, there must exist some $s \in T$ such that

$$U^3(a,1) \geq U^3(y(1),1) \text{ and } U^3(y(1),s) > U^3(a,s).$$

The first part of the expression implies that $y(1) = a, b$, or 0. None of these satisfy the second portion of the expression. Thus, in a complete information setting, the SCC F is not fully implementable.

The example has shown that a change in incentives, due to a change in information, can alter the credibility of agents and thus the class of implementable actions. In the complete information setting, agent 2 is helpless to prevent the deviation α', which he is able to prevent if information is incomplete. When information is complete, for agent 2 to credibly indicate that a deviation has occurred it is required that $U^2(a) \geq U^2[y(1)]$. Thus for agent 2 to be credible, she can't benefit from indicating that a deviation has occurred [claiming that state 2 is the true state and agents are deviating via α'] when one has not [the true state is 1]. The restriction for credibility is weaker when information is incomplete since agent 2 does not know the true state. Thus in this example, knowing more makes it possible for agent 2 to gain from falsely accusing the others of deviating and therefore ruins her credibility.

5. Discussion and Concluding Remarks.

One way in which the approach we have taken is not parallel to the work of Maskin (1977) is in that we have added assumptions (A1) and (A2). In order to make progress on the implementation problem with incomplete information these assumptions seem necessary for the following reasons. From the point of view of a planner, an incomplete information setting differs from a complete information setting in basic ways. For example, in a complete information setting, if only one agent is deviating from a given equilibrium play, the agent is identifiable, while in an incomplete information setting the agent may not be. [By identifiable, we mean that a mechanism can be constructed through which another agent can benefit from choosing an action which indicates that an agent has deviated, such that the agent only benefits from choosing this action if a deviation has indeed occurred.] In fact, in an incomplete information setting it may often be the case that a mechanism designer cannot even identify a subset of agents, some of whom may not be complying with a given equilibrium. For instance, consider a direct revelation mechanism for which s is announced such that $s \in B$ and $s \notin T$, where $B = \{s \in S \mid \forall i \; \exists t^i : s/t^i \in T\}$. In this case someone is lying, however, since for any i there exists t^i such that $s/t^i \in T$, it could be any agent or set of agents. Because of these sorts of possibilities, it is difficult to see a way to avoid the assumption (A1) in an incomplete information setting. The existence of the zero allocation allows the mechanism designer to deter deviations, when the deviating agent cannot be identified.

The assumption (A2) is satisfied trivially in any complete information setting. The reason for requiring it in an incomplete information setting is that the mechanism must be able to reward an agent who is indicating that a deviation has occurred. Since such an agent may not know what the true state is, it must be that the reward is beneficial on all of the agent's information set.

Our work in the incomplete information setting has brought us roughly to the point to which Maskin (1977) had brought us for the complete information setting, over a decade ago.[10] Recently, for the complete information setting, considerable progress has been made in characterizing the class of fully implementable social choice correspondences for stronger notions of equilibirum and weaker notions of implementation. [See Moore and Repullo (1986), Palfrey and Srivastava (1986), and Abreu and Sen (1987).] The stronger notions of equiblibrium and the weaker notions of implementation assure that almost all social choice correspondences are implementable. Palfrey and Srivastava (1987b) show the same is true for a private value setting. Whether similar results obtain in general incomplete information settings appears to be an open question.

The example we presented shows that there are important and intuitive differences between implementable allocations in complete and incomplete information settings. We have shown that there exists an allocation which is fully implementable with incomplete information, but not with complete information, and Palfrey and Srivastava (1985) have demonstrated social choice correspondences for which the converse is true. These examples give us some idea of the implications of different information structures, yet much remains to be understood in this area. A fruitful area for future research seems to be an application of the theory designed to identify how the information structure of an environment influences the outcomes of mechanisms. One possible use for such an application would be the characterization of the value of information in various settings.

---

[10] See also Williams (1984) and Saijo (1985).

References

ABREU, D. AND A. SEN [1987], "Virtual Implementation in Nash Equilibria," Department of Economics, Harvard.

ARROW, K. [1963], *Social Choice and Individual Values*, Cowles Foundation Monograph 12, Yale University Press: New Haven.

DASGUPTA, P., P. HAMMOND, AND E. MASKIN [1979], "The Implementation of Social Choice Rules: some General Results on Incentive Compatibility," *Review of Economic Studies*, Vol. 46, pp. 185–216.

HARRIS, M. AND R. TOWNSEND [1981], "Resource Allocation with Assymetric Information," *Econometrica*, Vol. 49, pp. 33–64.

HURWICZ, L. [1973], "The Design of Mechanisms for Resource Allocation," *American Economic Review*, Vol. 61, pp. 1–30.

JACKSON, M. [1988], "Private Information and Exchange: The Implications of Strategic Behavior," Ph.D. Dissertation: Graduate School of Business, Stanford University.

MASKIN, E. [1977], "Nash Equilibrium and Welfare Optimality," Department of Economics, MIT.

MASKIN, E. [1985], "The Theory of Implementation in Nash Equilibrium." In Ed.: L. Hurwicz, D. Schmeidler, and H. Sonnenschein, *Social Goals and Social Organization: Essays in Honor of Elisha A. Pazner*, Cambridge: Cambridge University Press.

MOORE, J. AND R. REPULLO [1986], "Subgame Perfect Implementation," Working Paper, London School of Economics.

MOOKHERJEE, D. AND S. REICHELSTEIN [1987], "Implementation via Augmented Revelation Mechanisms," Working Paper , Stanford Graduate School of Business.

MYERSON, R. [1979], "Incentive Compatibility and the Bargaining Problem," *Econometrica*, Vol. 47, pp. 61–74.

PALFREY, T., AND S. SRIVASTAVA [1985], "Implementation with Incomplete Information in Exchange Economies ," Carnegie Mellon University, Revised: 1987.

——————[1986], "Nash Implementation Using Undominated Strategies," Carnegie Mellon University.

——————[1987], "On Bayesian Implementable Allocations," *Review of Economic Studies*, Vol. 54, pp. 193–208.

——————[1987b], "Mechanism Design with Incomplete Information: A Solution to the Implementation Problem," Carnegie Mellon University.

POSTLEWAITE, A. AND D. SCHMEIDLER [1986], "Implementation in Differential Information Economies," *Journal of Economic Theory*, Vol. 39, pp. 14–33.

SAIJO, T. [1985], "Strategy Space Reduction in Maskin's Theorem: Sufficient Conditions for Nash Implementation," Cal. Tech..

WILLIAMS, S. [1984], "Sufficient Conditions for Nash Implementation," University of Minnesota.

APPENDIX

Proofs of Theorems 1, 2, and Remarks 1 and 2

PROOF OF THEOREMS 1 AND 2.

It was shown in Lemmas 1, 2, and 3 that a fully implementable SCC satisfies (SS), (BM), and (C). It is shown below that a SCC satisfying (C), (MNV), and (SS) is fully implementable. This establishes both theorems, since (BM) and (MNV) are equivalent, when (E) holds. This is accomplished by demonstrating a mechanism which fully implements a given SCC.

The mechanism which is described below is an 'augmented revelation mechanism'. [See Mookherjee and Reichelstein (1988). The mechanism described does not fit their definition precisely. Their definition requires that agents submit either their 'type' or an alternative message.] That is, the message space of each agent is a direct announcement of an agent's information ($s^i \in S^i$), coupled with an additional message. It is similar in design to the one used in Palfrey and Srivastava (1985). In this case, the additional message consists of: a state–contingent allocation, an integer, a fixed allocation, and either 0 or 1. The mechanism is designed so that an equilibrium involves agents announcing their true information set, or one consistent with (BM), and a state–contingent allocation in F. The remaining portions of the message space, are designed for use in the event of a deviation by some agent. In such an event, another agent signals that a deviation has occurred by announcing a high integer (or in some cases a 1), and is rewarded through the fixed allocation.

Let $M^i = S^i \times X \times \{0, 1, 2, \ldots\} \times \mathcal{A} \times \{0, 1\}$ and $B = \{s \in S \mid \forall i \; \exists t^i : s/t^i \in T\}$. For a given SCC, F, define the sets:

$$
\begin{aligned}
D_1 =& \{m \in M \mid m_1 = s, \; s \in B; \exists x \in F \text{ s.t. } m_2^i = x \; \forall i; \\
& \quad m_3^i = 0 \text{ for at least N} - 1 \text{ agents}, \; m_5^i = 0 \; \forall i\}, \\
D_2 =& \{m \in M \mid \exists k \text{ s.t. } m^k = (s^k, y, n^k, a^k, 1); \; \exists x \in F \text{ s.t. } m^i = (s^i, x, 0, a^i, 0) \; \forall i \neq k \\
& \quad x \; R^k[\pi^k(t^k)] \; y_{s^k} \forall t^k \text{ s.t. } s^{-k}/t^k \in T\}, \\
D_3 =& \{m \in M \mid m \notin D_2; \; \exists k \text{ s.t. } m^k = (s^k, y, n^k, a^k, 1); \\
& \quad \exists x \in F \text{ s.t. } m^i = (s^i, x, 0, a^i, 0) \; \forall i \neq k\}, \\
D_4 =& \{m \in M \mid m \notin D_1, \; \exists x \in F \; m^i = (s^i, x, 0, a^i, 0) \; \forall i\}, \\
D_5 =& \{m \in M \mid m \notin (D_1 \cup D_2 \cup D_3 \cup D_4), \;\; \exists k, \;\text{ and }\; \exists \tilde{m}^k, \;\text{ s.t. }\; m/\tilde{m}^k \in D_1 \\
& \quad \exists x \in F \;\; s.t. \;\; m^i = (s^i, x, 0, a^i, 0), \forall i \neq k; \}, \\
D_6 =& \{m \in M \mid m \notin (D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5)\}.
\end{aligned}
$$

Define the payoff function $g : M \to X$ by

$$
\begin{aligned}
g(m) &= x^0(s), \;\; m \in D_1, \\
g(m) &= y(s), \;\; m \in D_2, \\
g(m) &= 0, \qquad m \in D_3 \cup D_4 \cup D_5, \\
g(m) &= m_4^j, \;\; m \in D_6, \; j \in H, \; h = 1, \;\; \text{and} \\
g(m) &= 0, \qquad m \in D_6, \; h \neq 1,
\end{aligned}
$$

where $\bar{h} = \max_j \{m_3^j\}$, $\;H = \{j | m_3^j = \bar{h}\}$, and $h = \#H$.

The following claims are useful in the proof of Theorems 1 and 2.

CLAIM 1. *For any $x \in F$, σ, where $\sigma^i(s^i) = (s^i, x, 0, w^i(s^i), 0)$, is an equilibrium and $g(\sigma) = x^0$.*

CLAIM 2. *If σ is an equilibrium, then, for any $s \in T$: $\sigma(s) \in D_1$, $\sigma(s) \in D_2$ and $U^j(g[\sigma(s)], s) = U^j(w^j(s^j), s)$ $\forall j \neq k$, or $\sigma(s) \in D_6$ and $U^i(g[\sigma(s)], s) = U^i(w^i(s^i), s)$ $\forall i$.*

Claim 1 establishes that for each $x \in F$, there exists an equilibrium σ such that $g[\sigma(s)] = x(s)$ for all $s \in T$. Thus the first portion of the definition of full implementation is satisfied. Claim 2 will be used in establishing that the second portion of the definition is satisfied.

PROOF OF CLAIM 1.

Given an arbitrary $x \in F$ we consider σ defined by $\sigma^i(s^i) = (s^i, x, 0, w^i(s^i), 0)$. Since for all s, $\sigma(s) \in D_1 \cup D_4$, it follows that $g[\sigma(s)] = x^0(s)$ for all s. Next we verify that σ is an equilibrium by showing that there are no improving deviations.

Consider a deviation by some i at some $s^i \in S^i$. The deviation must be of the form $(\hat{s}^i, y, n, z, m)$.

If $m = 0$ and $y \neq x$ then the set of agents actions lies in D_5 which yields an allocation of 0 which cannot be improving.

If $m = 0$ and $y = x$ then we are in $D_1 \cup D_4 \cup D_5$, and so for any t such that $t^i = s^i$ the allocation is $[x^0]_{\hat{s}^i}(t)$. From (SS) we know that this is not improving.

If $m = 1$ then we are in $D_2 \cup D_3$. If we are in D_2 then the allocation is $y_{\hat{s}^i}$, while otherwise the allocation is 0. Let $z(t)$ be the allocation at t, where $t^i = s^i$. Then $z(t) = y_{\hat{s}^i}(t)$ when we are in D_2 and $z(t) = 0$ otherwise. It follows that $y_{\hat{s}^i}$ R^i $[\pi^i(s^i)]z$. If for any t we are in D_2 then x $R^i[\pi^i(s^i)]$ $y_{\hat{s}^i}$. Hence x $R^i[\pi^i(s^i)]$ z and the deviation is not improving. If we end up in D_3 then the allocation is 0 which is not improving.

We have proven Claim 1 by showing that there are no improving deviations.

PROOF OF CLAIM 2.

First, let us define $\widetilde{\sigma}^j$, a deviation from σ^j at $\bar{s}$ by agent j, by $\widetilde{\sigma}^j_1 = \sigma^j_1$, $\widetilde{\sigma}^j_2 = \sigma^j_2$,

$\tilde{\sigma}_5^j = \sigma_5^j$, and:

$$\begin{aligned}\tilde{\sigma}_3^j(\bar{s}^j) =& 1 + \max_{k,t}\{\sigma_3^k(t^k)\},\\ \tilde{\sigma}_3^j(t^j) =& \sigma_3^j(t^j), \quad \forall t^j \neq \bar{s}^j,\\ \tilde{\sigma}_4^j(\bar{s}^j) =& w^{\bar{s}^j},\\ \tilde{\sigma}_4^j(t^j) =& \sigma_4^j(t^j), \quad \forall t^j \neq \bar{s}^j.\end{aligned}$$

The following lemma is useful. Let $\tilde{\sigma} = \sigma/\tilde{\sigma}^j$.

LEMMA 4. *For a given set of strategies, in deviating to $\tilde{\sigma}^j$, agent j is at least as well off as using σ^j, for any s.*

PROOF: If $\sigma(s) \in D_1 \cup D_2 \cup D_3$, then either $\tilde{\sigma}(s) \in D_1 \cup D_2 \cup D_3$ and the allocation is unchanged, or $\tilde{\sigma}(s) \in D_6$ and the agent receives $w^j(s^j)$ and is at least as well off. If $\sigma(s) \in D_4 \cup D_5$, then the agent can be no worse off. If $\sigma(s) \in D_6$, then $\tilde{\sigma}(s) \in D_6$ and the agent receives $w^j(s^j)$ and is at least as well off. ∎

Consider an equilibrium σ, and suppose that Claim 2 does not hold. There are three (mutually exclusive) possibilities. For some $\bar{s} \in T$:

Case (1) $\sigma(\bar{s}) \in D_3$ or $\sigma(\bar{s}) \in D_2$ and $U^j(g[\sigma(s)], s) < U^j(w^j(s^j), s)$ for some $j \neq k$,

Case (2) $\sigma(\bar{s}) \in D_4 \cup D_5$, or

Case (3) $\sigma(\bar{s}) \in D_6$ and $U^j(g[\sigma(s)], s) < U^j(w^j(s^j), s)$ for some j.

We will show that in each case there is an agent j who is made better off by deviating to $\tilde{\sigma}^j$. We treat the three cases separately.

Case (1).

Let k be the agent defined in $D_2 \cup D_3$. There exists $j \neq k$ such that $U^j(g[\sigma(\bar{s})], \bar{s}) < U^j(w^j(\bar{s}^j), \bar{s})$ (recall that if $\sigma \in D_3$, then $g[\sigma(s)] = 0$). Let j deviate according to $\tilde{\sigma}$. Then $\tilde{\sigma}(\bar{s}) \in D_6$ and $g[\tilde{\sigma}(\bar{s})] = w^j(\bar{s}^j)$, which is a strict improvement for j at $\bar{s}$.

We have shown that $\tilde{\sigma}^j$ offers j a strict improvement over σ^j at $\bar{s} \in T$ and, by Lemma 4, $\tilde{\sigma}$ always does at least as well as σ^j. Since $q^j(\bar{s}) > 0$, we have contradicted the fact that σ is an equilibrium. This completes Case (1).

Case (2).

By the definition of $D_4 \cup D_5$, since $N \geq 3$, we can find j such that there does not exist m^j with $\sigma(\bar{s})/m^j \in D_1$. [In D_5 any $j \neq k$ will do, since it must be that $m_2^k = y \neq x$ or

$m_5^k = 1$, and so $\sigma(\bar{s})/m^j \notin D_1$, for all m^j. In D_4, since it must be that $\sigma(\bar{s})_1 \notin B$, there is some j such that $\sigma(\bar{s})_1/s^j \notin T$ for all s^j, and thus $\sigma(\bar{s})_1/s^j \notin B$ for all s^j.] If j deviates according to $\widetilde{\sigma}^j$, then $\widetilde{\sigma}(\bar{s}) \in D_6$ and $g(\bar{s}) = w^j(\bar{s}^j)$. Since $g[\sigma(\bar{s})] = 0$, $\widetilde{\sigma}^j$ offers a strict improvement for j at $\bar{s}$.

We have shown that $\widetilde{\sigma}^j$ offers j a strict improvement over σ^j at $\bar{s} \in T$, and by Lemma 4, $\widetilde{\sigma}$ always does as well as σ^j. Since $q^j(\bar{s}) > 0$, we have contradicted the fact that σ is an equilibrium. This completes Case (2).

Case (3).

Let j [The agent defined in Case (3)] deviate according to $\widetilde{\sigma}$. At $\bar{s}$, j now receives $w^j(\bar{s}^j)$ which is a strict improvement.

We have shown that $\widetilde{\sigma}^j$ offers j a strict improvement over σ^j at $\bar{s} \in T$, and by Lemma 4, $\widetilde{\sigma}^j$ always does as well as σ^j. Since $q^j(\bar{s}) > 0$, we have contradicted the fact that σ is an equilibrium. This completes Case (3).

We have proven Claim 2 by showing that none of Cases (1)–(3) are possible for an equilibrium σ.

To prove Theorems 1 and 2, it remains to be shown that for each equilibrium, as discussed in Claim 2, there exists a $z \in F$ such that $z(s) = g[\sigma(s)]$ for all $s \in T$.

Define $C = \{s : \sigma(s) \in D_1 \text{ and } \sigma_3^i(s) = 0\ \forall i\}$.

From Claim 2, we know that, for any equilibrium σ and $s \in T$, if $\sigma(s) \notin D_1$, then (NVH) applies. For states $s \in T$ such that $s \notin C$ and $\sigma(s) \in D_1$, it must be that $\sigma_3^i(s) \neq 0$ for some i. In this case, it must be that $U^j(g[\sigma(s)], s) = U^j(w^j(s^j), s)$ for all $j \neq i$ [or else some agent j could deviate according to $\widetilde{\sigma}$ as defined in Claim 2 and be made better off]. Therefore, in this case (NVH) applies as well, and so (NVH) is satisfied at all $s \in T$ such that $s \notin C$.

Now consider the set C. Since $\sigma(s) \in D_1$ for all $s \in C$, there exists α such that for each $s \in C$ there exists $x_s \in C$ such that $g[\sigma(t)] = x_s^0[\alpha(t)]\ \forall t \in C \cap [\cup_i \pi^i(s^i)]$. Suppose that there does not exist a $\widehat{z} \in F$ such that $\widehat{z}(s) = g[\sigma(s)]$ for all $s \in T$. From (MNV) we know that there exist i, y, and $s \in C$ such that

$$y \circ \alpha\ P^i[\pi^i(s^i) \cap C]\ g(\sigma), \text{ while}$$

$$x_s \ R^i[\pi^i(t^i)] \ y_{\alpha^i(s^i)} \ \forall t^i : \alpha(s)/t^i \in T.$$

Therefore i is better off submitting $[\alpha^i(s^i), y, 1 + \max_{k.t}\{\sigma_3^k(t^k)\}, w^i(s^i), 1]$ whenever s^i is observed. This is shown as follows: The deviation puts the action in D_2 for all $s \in C \cap \pi^i(s^i)$ and the outcome is $y \circ \alpha$ which is strictly preferred by i to $g(\sigma)$ on the set $C \cap \pi^i(s^i)$. We now show that for any other $t \in \pi^i(s^i)$ $(t \notin C)$, the deviation does at least as well for i. If $\sigma(t)$ was in D_1, since $t \notin C$ there was some other agent k with $\sigma^k(t)_3 > 0$, and so the new action is in D_6 and the outcome is $w^i(s^i)$. If $\sigma(t)$ was in D_2, then there was some other agent k with $\sigma^k(t)_5 = 1$, and so the new action is in D_6 and the outcome is $w^i(s^i)$. If $\sigma(t)$ was in D_6 and $g[\sigma(t)] \neq 0$ then there was some other agent k with $\sigma^k(t)_3 > 0$, and so the new action is still in D_6 and the outcome is $w^i(s^i)$; while if $g[\sigma(t)] = 0$ then the new action does at least as well.

We have contradicted the fact that σ is an equilibrium. Hence, our supposition was wrong and there exists a $z \in F$, such that $z(s) = g[\sigma(s)]$ for all $s \in T$.

We have shown that if σ is an equilibrium, then there is a $z \in F$ such that $z(s) = g[\sigma(s)]$ for all $s \in T$. This, coupled with Claim 1, establishes Theorems 1 and 2.

PROOFS OF REMARKS 1 AND 2.

Remark 1 follows directly from Claim 1, with the remark that only (SS) is needed for the proof of Claim 1.

Remark 2 follows directly from the proof of Theorems 1 and 2, with the remark that (SS) is only needed in the proof of Claim 1, and so only (C) and (MNV) are used in the remainder of the proof. Given (E), (MNV) and (BM) are equivalent.