VIRTUAL IMPLEMENTATION IN ITERATIVELY UNDOMINATED STRATEGIES:

INCOMPLETE INFORMATION*

Dilip Abreu

Princeton University

and

Hitoshi Matsushima

The University of Tsukuba

September 1990
Revised, August 1992

* The authors wish to thank The Hoover Institution, The Graduate School of Business and The Department of Engineering-Economic Systems at Stanford University, The Japan Society for the Promotion of Science, The National Science Foundation and The Sloan Foundation for their support.
ABSTRACT

extend our results for complete information to general Bayesian environments. A social choice function now maps from a finite set of profiles of signals or types to lotteries over alternatives. Self-selection is an obvious necessary condition for virtual implementation under any solution concept. We derive an additional necessary condition which we term measurability. Measurability requires that the social choice function be measurable with respect to a particular partition of each player's signal space. These partitions are derived from players' state-dependent utility functions and their conditional probability distributions over other players' types. The condition will often be trivially satisfied; the relevant players' partitions will be the finest possible, each element of each player's partition containing a single signal. Under weak domain restrictions we show that any social choice function which satisfies self-selection and measurability is virtually implementable in iteratively undominated strategies. This result is essentially complete, and as permissive as one might hope. It applies to an arbitrary number of players and in particular embraces the two player complete information case.

preceding literature on Bayesian (and Nash) implementation is seriously undermined by its reliance on "integer" games and/or by the arbitrary exclusion of mixed strategies. The latter limitation is severe for a research program whose primary objective is accounting for all equilibrium possibilities. The mechanism we construct here is finite, and yields, as a corollary, virtual implementation in (pure and) mixed Bayesian Nash equilibrium. This equilibrium is unique, pure and strict
1. INTRODUCTION

This paper considers implementation in Bayesian environments with two or more players. It extends our results for complete information in Abreu and Matsushima 1996 to this more general setting. In terms of motivation and general background the complete and incomplete information problems have much in common. To avoid repetition, this introduction, and more generally the paper, is written to be read in conjunction with our companion piece on complete information.

The setting is the following. There are a collection of alternatives a group of players, each of whom receives a private signal. The set of possible signals or types is finite. Every player has a distribution over players' signals, conditional on the signal she receives. (In the special case of complete information each player receives the same signal.

Players' preferences over lotteries are, in general, determined by the entire profile of players' signals. A social choice function maps from the space of possible signal profiles to a (fixed) set of possible outcomes, indicating the signal-contingent outcome the principal or planner wishes to impose. An essential feature of our formulation is that the set of outcomes is taken to be the space of (simple) lotteries over some (arbitrary) set of pure alternatives. We will assume that players' preferences over lotteries satisfy the expected utility hypothesis.

A game form or mechanism consists of a message space for each player and an outcome function which specifies an outcome for each message profile. A social choice function is implementable via the iterative elimination of
strictly dominated strategies if there exists a game form for which the iterative elimination of strictly dominated strategies leads to a unique iteratively undominated strategy profile with the required outcome for all (possible) signal profiles.

A social choice function satisfies self-selection if truthful reporting is a best response given that other players report their signals truthfully. That is, truthful reporting is a Bayesian Nash equilibrium in the standard revelation mechanism in which players only report their signals and the outcome function is simply the social choice function itself. Self-selection is a necessary condition for implementation (virtual or exact) using any solution concept. We derive a second necessary condition which is new to the literature. We term this condition measurability. It requires that the social choice function be measurable with respect to a particular partition of each player's signal space. The relevant partitions are determined by players' state-dependent utility functions, and their conditional distributions over other players' signals. Measurability will, of course, be trivially satisfied if the partition is the finest possible, each element of the partition being a singleton. This is in fact the case in complete information environments, and we indicate plausible conditions under which will be true in general incomplete information environments also. Significantly, measurability is a necessary condition for implementation (virtual and exact) in Bayesian Nash equilibrium also, when the implementing form satisfies a natural regularity condition. Hence in situations in which the measurability condition is non-trivial it cannot be evaded by weakening the solution concept to Bayesian Nash equilibrium.
Under weak domain restrictions on preferences we show that any social choice function which satisfies self-selection and measurability is virtually implementable via the iterative elimination of strictly dominated strategies. This characterization is essentially complete. Given our domain restrictions, self-selection and measurability are necessary and sufficient for virtual implementation. Since these conditions are also necessary for implementation in Bayesian Nash equilibrium it shows that (modulo virtualness) there is no advantage in weakening the solution concept beyond the iterative removal of strictly dominated strategies.

We are aware of no characterizations involving iterative dominance arguments in general Bayesian environments. There is however a fairly large body of work on Bayesian Nash implementation in incomplete information settings starting with Postlewaite and Schmeidler (1986). Subsequent contributions include Palfrey and Srivastava (1989a, 1989b), Mookherjee Reichelstein (1989), Jackson (1990, 1991), Matsushima (1990, 1991), and others. For further references see the survey by Palfrey (1990).

This literature shows that self-selection and a Bayesian version of the monotonicity condition are necessary for Bayesian Nash implementation. As in the complete information case there is no presumption that monotonicity is a non-trivial condition in general. All the drawbacks of the Nash literature carry over to the Bayesian Nash case. These include the use of integer games and the arbitrary exclusion of mixed strategies. See Jackson (1990b), and our companion paper for a critique. In contrast our results are permissive, obtained for a much weaker solution concept, and use finite mechanisms. The unique iteratively
undominated strategy profile is a strict Bayesian Nash equilibrium. Moreover, it is a unique equilibrium. Furthermore, our construction applies to an arbitrary number of players, and consequently provides a unified treatment of the two and more than two player case. The two player case has not been considered in the Bayesian implementation literature and has needed new methods of proof in the complete information case.

This paper is organized as follows. Section 2 develops notations. Section 3 develops a special case. Section 4 discusses measurability and Section 5 presents the theorem. Section 6 shows how the theorem applies to complete information and the two player case. Section 7 concludes.
2. PRELIMINARIES

Let $N = \{1,\ldots,n\}$ denote the set of individuals (agents, players), and $S_i$ be the finite set of types or signals of individual $i$. A profile of signals defines a state $s = (s_1, \ldots, s_n)$ and $S = \times_{i \in N} S_i$ denotes the set of possible states. Let $p_i(s_{-i}; s_i)$ denote agent $i$'s conditional probability that other agents receive the profile of signals $s_{-i}$ when she receives the signal $s_i$. As in our companion paper, $A$ denotes the set of simple lotteries over some arbitrary set of pure alternatives. A pair of lotteries $a, b \in A$ is $\varepsilon$-close if the distance between them is at most $\varepsilon$ in the usual Euclidean metric $\sum_{\tau \in \Gamma} (a(\tau) - b(\tau))^2 \leq \varepsilon$, where $\Gamma$ is the support of lotteries $a$ and $b$ and $a(\tau)$ is the probability of pure alternative $\tau$ in the simple lottery $a$.

Player $i$'s state-dependent von Neumann-Morgenstern utility function is denoted $u_i : A \times S \rightarrow R$. It is linear in its first argument. Note that $u_i$ in general depends on all players' signals.

A social choice function $x : S \rightarrow A$ maps from states to lotteries. We will write $x = \lambda y + (1 - \lambda)z$ if $x(s) = \lambda y(s) + (1 - \lambda)z(s)$ for all $s \in S$. The social choice functions $x$ and $y$ are $\varepsilon$-close if for all $s \in S$, the lotteries $x(s)$ and $y(s)$ are $\varepsilon$-close. Let

$$U_i(x, s_i) = \sum_{s_{-i} \in S_{-i}} u_i(x(s), s)p_i(s_{-i}; s_i)$$

denote individual $i$'s conditional expected utility from a social choice function $x$ when she receives the signal $s_i$. A mechanism or a game form $G$ is
an \((n+1)\)-tuple \((M_1, \ldots, M_n; g)\), where \(M_i\) is a message space for agent \(i\), \(M = M_1 \times \cdots \times M_n\), and \(g: M \rightarrow A\) is an outcome function. Our constructions only use finite \(M_i\)'s. Let \(a_i: S_i \rightarrow M_i\) denote a (pure) strategy for agent \(i\) and \(s_i\) the set of pure strategies. Let

\[
v_i(G, a, s_i) = \sum_{s_{-i} \in S_{-i}} u_i(g(a(s)), s)p_i(s_{-i}|s_i)
\]

denote agent \(i\)'s conditional expected utility from a mechanism \(G\) under the strategy profile \(a\), when \(s_{-i}\) receives the signal \(s_i\). Note that a player's strategy specifies a message for each of the possible types.

Fix a game form \(G = (M, g)\) arbitrarily. Let \(H_1\) be a subset of \(E_1\). A strategy \(a_i \in H_1\) is strictly dominated for player \(i\) with respect to \(H = \times_{j \in N} H_j\) if there exist \(a'_i \in H_1\) and \(s_i \in S_i\) such that for every \(a_{-i} \in H_{-i}\),

\[
v_i(G, a/a'_i, s_i) > v_i(G, a, s_i).
\]

Note that from the point of view of a player \(i\) who receives signal \(s_i\) the domination is strict with respect to all possible strategies of other players. Let \(Q_1(H)\) denote the set of all undominated strategies for agent \(i\) with respect to \(H\). Let \(Q(H) = \times_{i \in N} Q_i(H)\). Let \(Q^k(\Sigma) = \times_{i \in N} Q^k_i(\Sigma)\), where \(Q^k_i(\Sigma) = Q_i(Q^{k-1}_i(\Sigma))\). For simplicity, we write \(Q^k\) for \(Q^k(\Sigma)\) etc.

Let \(Q^k\) denote the intersection of \(Q^k, k = 0, 1, \ldots\). A strategy profile \(a \in \Sigma\) is iteratively undominated if \(a \in Q^k\). Since \(\Sigma\) is finite, there exists \(k\)
such that $Q^k = Q^k$ for all $k \geq k$. The mechanism $G = (M, g)$ exactly implements a social choice function $x$ in iteratively undominated strategies if and only if $Q^*$ is a singleton and $g(\delta^*(s)) = x(s)$ for all $s \in S$, where $Q^* = \{\delta^*\}$.

As in our earlier paper, the order of elimination of strictly dominated strategies is irrelevant. If $\delta^*$ is the unique strategy which survives the iterative elimination of strictly dominated pure strategies, it will also be the unique strategy which survives the iterative elimination of strictly dominated (pure and) mixed strategies. Furthermore, $\delta^*$ is a unique mixed Bayesian Nash equilibrium, and is a strict equilibrium for each type of every player.

For every $i \in N$, every $s_i \in S_i$ and every $s' \in S_i$, let

$$V_i(x, s_i, s') = \sum_{s_{-i} \in S_{-i}} u_i(x(s/s'), s)p_i(s_{-i}; s_i)$$

denote the expected utility of the direct mechanism $(S, x)$ for player $i$ conditional on $s_i$ when he announces $s'_i$ and the other players make truthful announcements. Note that

$$V_i(x, s_i, s_i) = U_i(x, s_i)$$

**DEFINITION:** A social choice function $x$ satisfies self-selection if for every $i \in N$, every $s_i \in S_i$ and every $s'_i \in S_i/(s_i)$,

$$U_i(x, s_i) \geq V_i(x, s_i, s'_i).$$

A social choice function $x$ satisfies strict self-selection if these inequalities strictly hold.
Self-selection means that truth-telling is a Bayesian Nash equilibrium in the direct mechanism \((S,x)\). Self-selection is a well-known (and immediate) necessary condition for implementation in Bayesian Nash equilibrium, or in any refinement of Bayesian Nash equilibrium.
3. A SPECIAL CASE

Before proceeding to our general characterization we consider a special which we hope will serve as an easy introduction to the incomplete information setting. In this section we assume

that small side-payments are possible and that these affect utility additively. Specifically, we denote by $t_i \in [-\epsilon, \epsilon]$ the side-payment to player $i$ and suppose that her total utility is $u_i(a, s) + t_i$.

private values: Player $i$'s preferences over lotteries depend only on her own type. Since $u_i$ is independent of $s_{-i}$, we simply write $u_i(a, s_i)$ instead of $u_i(a, s_i, s_{-i})$.

that for every $i \in N$, and $s_i, s'_i$, there exist $a, a' \in A$ such that $u_i(a, s_i) > u_i(a', s_i)$. That is, for each type of player $i$, we exclude the possibility of universal indifference across lotteries that distinct types of player $i$ have distinct orderings over lotteries.

Assumption (4) is not innocuous. In particular it excludes the complete information case. A player's signal represents what she knows, and in the complete information setting this is the entire profile of players'
preferences. Hence two distinct signals for player 1 may represent differences in other players' preference orderings and not her own. See Section 4 for a further discussion of this point.

The preceding assumptions (specifically (2)-(4) imply that for every $i \in N$ there exists a function $f_i : S_i \mapsto A$ such that for every $s_i \in S_i$ and every $s'_i \in S_i / s_i$

$$u_i(f_i(s_i), s_i) > u_i(f_i(s'_i), s_i).$$

This straightforward result corresponds to the Lemma in Aziz and Matsushima (1992), and may be proved analogously. We may think of $f_i$ as a social choice function which is independent of $s_i$. It gives player i's "dictatorial" choice within the set $\{ a : a = f_i(s_i) \text{ for some } s_i \in S_i \}$. By construction this choice varies with $s_i$.

We now construct a mechanism as follows. Every player $i$ makes $(K+1)$ simultaneous announcements, each of which is of his own signal $M_i = M_i^0 \times M_i^1$ $M_i^K \times S_i \times S_i$. Recall that $S_i$ is finite so that this construction will yield a finite mechanism. Denote

$$m_i = (m_i^0, \ldots, m_i^K) \in M_i^0 \times M_i^1$$

$$m = (m^0_i, \ldots, m^K_i) \in M_i^0 \times M_i^1$$

$$m = (m^h_1)_{i \in N} \in M_i^h$$

Given a social choice function $x$, for any profile of player messages $m$, the lottery chosen is
\[ g(m) = \frac{\varepsilon}{n} \sum_{i \in N} f_i(m_1^0) + \frac{1}{K} \sum_{h=1}^{K} \left( \frac{\varepsilon}{n} \sum_{i \in N} f_i(m_i^h) + 1 - \varepsilon - \varepsilon^2 \right) x(m^h) \]

where \( \varepsilon \) is small and strictly positive. Equivalently,

\[ g(m) = \frac{\varepsilon}{n} \sum_{i \in N} f_i(m_1^0) + \frac{1-\varepsilon}{K} \sum_{h=1}^{K} x'(m^h), \]

where

\[ x'(m^h) = \frac{\alpha}{n} \sum_{i=1}^{n} f_i(m_i^h) + (1-\alpha) x(m^h), \]

and

\[ \frac{\varepsilon^2}{1-\varepsilon} \]

that given that \( x \) satisfies self-selection, \( x' \) satisfies strict self-selection. This is because of the addition of the \( f_i \) terms. Of course, for small \( \varepsilon \), \( x' \) is "close" to \( x \).

In the above mechanism, player \( i \)'s zero-th announcement affects the outcome with probability \( \frac{\varepsilon}{n} \). With this probability the outcome is \( f_i(m_1^0) \), and the lottery depends only on player \( i \)'s zero-th announcement. On the other hand, player \( i \)'s \( h \)-th announcement \( (h \geq 1) \) affects the outcome with probability \( \frac{1-\varepsilon}{K} \) via the term \( x'(m_1^h, m_{-1}^h) \) which does depend on other players' h-th announcements.

In addition to this lottery players receive small fines according to the following rules. The first deviant from \( i \)'s own zero-th announcement is fined \( n \). That is, player \( i \) is fined \( t_1 : M \rightarrow R \) where
\[ t_i(m) = -n \quad \text{if player } i \text{ is the first deviant,} \]

i.e., for some \( h, m^h_i = m_i^0 \) and \( m^h \)

\[ m^0 \text{ for all } h' < h \]

\[ t_i(m) = 0 \quad \text{otherwise.} \]

Choose \( n \) sufficiently small, such that for every \( i \in \mathbb{N}, \]

\[ \frac{\varepsilon}{n} \left( u_i(f_i(s_i), s_i) + u_i(f_i(s'_i), s_i) \right) \quad \text{for all } s_i \in S_i \text{ and all } \]

\[ s'_i \in S_i / (s_i). \]

Note that \( n \) is smaller than the "direct" expected utility loss from a zero-

th message \( m_i^0 \rightarrow s_i \).

Let \( \sigma \) be an iteratively undominated strategy profile. Recall that \( q_i : S_i \rightarrow M_i \).

Denote strategies for players by

\[ \sigma_i = (\sigma_i^0, \ldots, \sigma_i^K), \quad \sigma_i^h : S_i \rightarrow M_i^h, \]

\[ \sigma = (\sigma_1^0, \ldots, \sigma_i^K), \quad \sigma : S \rightarrow M^h \]

Since \( m_i^0 \) affects player \( i \)'s utility only through \( f_i \) and \( t_i \), it follows

directly from the definition of \( n \) that if \( \sigma_i \) is iteratively undominated, then

\[ \sigma_i^0(s_i) = s_i. \]

We will now show that if for every \( i \in \mathbb{N}, \) every iteratively undominated

strategy profile \( \sigma_i \) and every \( s_i \in S_i, \quad \sigma_i^h(s_i) = s_i \) for all \( h \in (0, \ldots, K), \) then
$o_1^{k+1}(s_1) = s_1$ for all $i \in N$ and all $s_1 \in S_1$

The following inductive step completes the argument: for any (small $\varepsilon$)

0. the mechanism yields a unique iteratively undominated profile $\sigma$ with

$k\sigma_i(s_1) = s_1$ for all $s_1 \in S_1$, all $i \in N$ and all $k = 0, ..., K$. The resultant outcome is

$$(1 - \varepsilon - \varepsilon^2)x(s) + \sum_{i \in N} \left( \frac{\varepsilon}{n} + \frac{\varepsilon^2}{n} \right) f_i(s_1),$$

and there are no fines.

For the inductive step to go through we must choose $K$ sufficiently large, such that for every $i \in N$ every $s_1 \in S_1$ and every $s' \in S$

$$n > \frac{1 - \varepsilon}{K} \{u_1(x'(s'/s_1), s_1) - u_1(x'(s), s_1)\}.$$ Then the fine $n$ is larger than $\frac{1 - \varepsilon}{K}$ times the maximal "direct" utility gain from changing the outcome.

Suppose that $o_1^{k+1}(s_1) \neq s_1$ for some $i \in N$ and some $s_1 \in S_1$. Define $o_1$ such that

$$o_1 = o_1 \text{ for } h \neq k+1, \quad o_1^{k+1}(s'_1) = o_1^{k+1}(s_1) \text{ for all } s'_1 \in S_1(s_1)$$

and $o_1^{k+1}(s_1) = s_1$.

Under the inductive hypothesis above if $o_j^{k+1}(s_j) = s_j$ for all $j \in N \setminus \{i\}$ and all $s_j \in S_j$, then by strict self-selection $o_1$ yields higher payoff than $o_1$. 

even if player i is the first to deviate given \((a_1, a_{-1})\). On the other hand, if for some \(j \in N(1)\), \(a_j(s_j) \triangleright s_j\) for all \(s_j \in S_j\), then (by the choice of \(K\) above) \(a_1\), by saving the fine \(n\), yields higher payoff than \(a_1\). Thus \(a_1\) dominates \(a_1\) for player i of type \(s_1\). In fact this argument is not complete in that some types of player \(j\) may announce \(a'_j(s_j) \triangleright s_j\), while others make truthful announcements. This possibility is allowed for in the detailed arguments we present below. Assumption (2) and (4) are the non-innocuous simplifying assumptions used here. These assumptions are relaxed in the next two sections which also provide formal proofs.
4. MEASURABILITY

This section introduces the **measurability condition** and discusses issues of computation. Establishes that measurability is necessary for implementation in Bayesian Nash equilibrium also, and finally provides a simple sufficient condition under which measurability is automatically satisfied by any \( \mathcal{S} \).

4.1 THE CONDITION

As indicated in the introduction, the measurability condition requires that the social choice function be measurable with respect to a particular partition of each player's signal space. In order to clarify exactly what this condition means it is useful to start with a little notation.

Denote by \( \psi_i \) a partition of \( S_i \), where \( \psi_i \) is a generic element of \( \psi_i \) and \( \psi_i(s_i) \) is the element of \( \psi_i \) which includes \( s_i \). Let \( \psi = \times_i \psi_i \) and \( \psi_i \) _\( \times N \).

**DEFINITION:** A social choice function \( x \) is **measurable** with respect to \( \psi \) if for every \( i \in \mathbb{N} \), every \( s_{-i} \in S_{-i} \) and \( s_i \in S_i \),

\[
x(s_i) = x(s_i/s_{-i}' \text{ for all } s_{-i} \in S_{-i} \text{ whenever } \psi_i(s_i) = \psi_i(s_{i}')
\]

A social choice function \( x \) is **strictly measurable** with respect to \( \psi \) if for every \( i \in \mathbb{N} \), every \( s_{-i} \in S_{-i} \) and \( s_{i}' \in S_i \),
\[ x(s) = x(s/s'_1) \text{ for all } s_{-1} \in S_{-1} \text{ if and only if } \gamma_1(s'_1) = \gamma_1(s'_1). \]

Moreover, a social choice function \( x \) is measurable with respect to a social choice function \( y \) if for any \( i \in N \) and any \( s_1, s'_1 \in S_1 \),

\[ x(s) = x(s/s'_1) \text{ for all } s_{-1} \in S_{-1} \text{ whenever } y(s) = y(s/s'_1) \text{ for all } s_{-1} \in S_{-1}. \]

Measurability of \( x \) with respect to \( y \) implies that for any player \( i \), \( x \) does not distinguish between any pair of signals in the same element of the partition \( \psi_1 \).

**DEFINITION:** A strategy \( \sigma_i \) for player \( i \) is measurable with respect to \( \psi_i \) if for every \( s'_1 \in S_1 \) and every

\[ \sigma_i(s'_1) = \sigma_i(s'_1) \text{ whenever } \gamma_i(s'_1) = \gamma_i(s'_1). \]

A strategy profile \( \sigma \) is measurable with respect to \( \psi \) if for every \( i \in N \), \( \sigma_i \) is measurable with respect to \( \psi_i \).

Consider player \( i \), and \( Y^i \), the set of social choice functions which depend only on other players' signals. For any signal \( s_i \) player \( i \) receives, she can rank elements of \( Y^i \) in terms of conditional expected utility. In any game form, and for any tuple of strategies \( \sigma_{-i} \) of other players, a message
m_i by player i generates a social choice function in Y_i (y_i such that

= g(m_i, \theta^{-1}(s_{-i})

If the signals s_i and s'_i generate the same posterior ordering over Y_i then, in any game form, player i's optimal strategies when he is type s_i are identical to his optimal strategies when he is type s'_i. Player i could not possibly have a strict incentive to send a distinct message for each of these signals. Hence if x is implementable it must be that x(s_i, s_{-i}) = x(s'_i, s_{-i}) for all s_{-i} \in S_{-i}. That is, x must be measurable with respect to some partition of S_i with s_i and s'_i in the same cell of this partition.

Partition generated by posterior orderings over Y_i is the finest possible partition with respect to which an implementable social choice function must be measurable. In general the relevant partition might be coarser. This is because in distinguishing between any pair of signals s_i and s'_i we are using the richest possible set of social choice functions Y_i which depends only on other players' signals. The same analysis applied to players j \neq i may yield non-trivial partitions \psi_j of the S_j's. Then in order to discriminate between player i's signals we may only use a set of social choice functions which are measurable with respect to \psi_{-i}. Proceeding in this way we iteratively obtain coarser and coarser partitions. As will become clear the procedure described below starts at the "other end" and iteratively yields finer and finer partitions.
every $i \in \mathbb{N}$, $s_i \in S_i$, $s' \in S_i$ and $(n-1)$-tuple of partitions $\psi_1$. $s_i$ is equivalent to $s'_i$ with respect to $\psi_1$ if for every $x$ and every $y$ which are measurable with respect to $(S_i \times \psi_1$)

$$U_i(x,s_i) \geq U_i(y,s'_i)$$

if and only if $U_i(x,s'_i) \geq U_i(y,s_i')$

Let $\mathfrak{P}(S_1, \psi_1$ be the set of all elements of $S_1$ equivalent to $s_i$ with respect to $\psi_1$, and let

$$R^1(\psi_1) = \mathfrak{P}(S_1, \psi_1; s_i \in S_i$$

We define an infinite sequence of $n$-tuples of partitions, $(\psi^h)_{h=0}^\infty$, $\psi^h = \psi_1 \times$ in the following way. For every $i \in \mathbb{N}$,

$$\psi_1^0 = \{S_i\},$$

and recursively, for every $i \in \mathbb{N}$ and every $h = 1, 2, \ldots$

$$\psi_1^h = R^1(\psi_1^{h-1})$$

Note that for every $h = 0, 1, \ldots, \psi_1^{h+1}$ is the same as, or finer than, $\psi_1^h$.

Since $S_i$ is finite, there exists a positive integer $L$ such that for every $h \geq L$, $\psi_1^h = \psi_1^L$. We denote $\psi = \psi^L$. We will argue that a necessary condition for a social choice function to be virtually implementable is that it be measurable with respect to $\psi$

**DEFINITION:** A social choice function satisfies the measurability condition if it is measurable with respect to $\psi$.
Consider a game form \( G = (M_1, \ldots, M_n; g) \) which exactly implements a social choice function \( \gamma \) in iteratively strictly undominated strategies and let \( Q^1, Q^k \), the sets of iteratively undominated strategies at the \( k \)-th round of iterative removal, etcetera, be defined as in Section 2.

Consider an arbitrary constant strategy profile \( \sigma[0] \in Q^0 \) (that is, is measurable with respect to \( \times \{ S_1 \}_{i \in N} \). By the definition of \( \psi^1 \), it follows that for every \( i \in N \), there exists \( \sigma[1] \in S_1 \) which is a best response to \( \sigma[0] \) and is measurable with respect to \( \psi^1 \). Hence, \( \sigma[1] \) is not strictly dominated for player \( i \) with respect to \( E \), that is, \( \sigma[1] \in Q^1 \).

Fix \( k = 2, 3, \ldots \) arbitrarily, and suppose that there exists a strategy profile \( \sigma[k-1] \in Q^{k-1} \) which is measurable with respect to \( \psi^{k-1} \). Then it is easy to see that for every \( i \in N \), there exists \( \sigma[k] \in S_1 \) which is a best response to \( \sigma[k-1] \) and is measurable with respect to \( \psi^k \). Since \( \sigma[k] \) is a best response it is not strictly dominated for player \( i \) with respect to \( E \), \( Q_{i-1}^k \). As we are eliminating strictly dominated strategies, \( Q_i^k = Q(S_i \times Q_{i-1}^{k-1}) \). Hence for all \( k = 0, 1, \ldots \), there exists \( \sigma[k] \in Q^k \) which is measurable with respect to \( \psi^k \).

Let \( \sigma^* \) be the unique iteratively undominated strategy profile in the implementing game form \( G \). Then the preceding argument implies that \( \sigma^* \) is
measurable with respect to $\psi^*$. It follows that $\gamma = g \cdot \sigma^*$ is measurable with respect to $\psi^*$ also, i.e., satisfies the measurability condition.

Now suppose that $\gamma^m_{m=1}\infty$ is a sequence of social choice functions such that for all $m = m, m+1$, $x$ is measurable with respect to $\gamma^m$ (and vice-versa). Thus if $x$ is virtually implementable there exists $\gamma$ which is measurable with respect to $x$ and exactly implementable. Hence we have proved

**PROPOSITION 1** If a social choice function is virtually implementable in iteratively undominated strategies, then it satisfies the measurability condition.

### 4.2. COMPUTING $\psi^*$

There are simple sufficient conditions under which no computation is necessary. The partition is simply the finest possible partition $\psi_1$, each element of $\psi_1$ containing only a single element. See Section 4.4. When computation is required and the set of pure alternatives $\Gamma$ is finite, $\psi^*$ may be determined via a finite procedure as follows. Define

$$x^D = \{x; x(s) \text{ is a degenerate lottery for all } s \in S\}$$
If $r$ is finite so also

The following definition of equivalence of signals is implied by and implies the definition given earlier: $s_1$ is equivalent to $s_1'$ with respect to $\mathcal{F}_1$ if there exist $\alpha$ and $\beta > 0$ such that for any $x \in X$ which is measurable with respect to $(S_1, \mathcal{F}_1, \mathcal{F}_1)$,

$$U_1(x, s_1) = \alpha + \beta U_1(x, s_1').$$

The equivalence of the two definitions follows from the fact that we need only consider social choice functions which map to pure outcomes and check the condition above.

4.3. MEASURABILITY AND BAYESIAN NASH IMPLEMENTATION

We argue here that measurability is a necessary condition for implementation in Bayesian Nash equilibrium. When implementing game forms satisfying a natural regularity condition, thus measurability is a weak condition; requiring implementation in iteratively undominated strategies does not entail requirements beyond these needed for Bayesian implementation.

For any mechanism $G = (M, g)$, let $\text{BN}(G)$ be the set of Bayesian Nash equilibria of $G$. A social choice function $x$ is implementable in Bayesian Nash equilibrium by the game form $G$ if $\text{BN}(G) \neq \emptyset$ and for all $\sigma \in \text{BN}(G)$

$$g(\sigma(s)) = x(s) \text{ for all } s \in S.$$
For every \( i \in \mathbb{N} \) and every partition \( \psi_i \), let \( \Sigma_i(\psi_i) \) denote the set of strategies of player \( i \) which are measurable with respect to \( \psi_i \). The profile \( \sigma \times \prod_{i \in \mathbb{N}} \Sigma_i(\psi_i) \) is a pseudo-Bayesian Nash equilibrium with respect to \( \psi \) in \( G \) if for all \( i \in \mathbb{N} \) and all \( \psi_i \), there exists some \( s_i \in \psi_i \) such that
\[
\nu_i(G, \sigma, s_i) \geq \nu_i(G, \sigma/\sigma_i', s_i) \quad \text{for all } \sigma_i' \in \Sigma_i
\]
We will say that \( G \) is regular if for any \( \psi \), a pseudo-Bayesian Nash equilibrium exists. Regularity is a minimal requirement, and will be satisfied, for instance, under any of the standard conditions used to establish the existence of a Bayesian Nash equilibrium via a fixed point argument.

PROPOSITION 2: If a social choice function is virtually implementable in Bayesian Nash equilibrium by a regular game form, then it satisfies the measurability condition.

PROOF: As in the earlier proof, it suffices to establish the result for exactly implementable social choice functions. Let \( G = (M, g) \) exactly implement a social choice function \( \chi \), and let \( \sigma \times \prod_{i \in \mathbb{N}} \Sigma_i(\psi_i) \) be a pseudo-Bayesian Nash equilibrium with respect to \( \psi^* \). If \( m_i = \sigma_i(s_i) \) is a best response for player \( i \) with signal \( s_i \) then \( m_i \) is also a best response for any \( s_i' \in \Sigma_i(s_i, \psi_i) \) where \( \eta_i(\psi_i) = \bigcup_{i \in \mathbb{N}} \sigma_i(t_{i, \psi_i}) \cap \Sigma_i(t_i, \psi_i) \).
follows that any pseudo-equilibrium $\sigma$ which is measurable with respect to $\psi^*$ is in fact a regular equilibrium. Since $x = g\cdot \sigma$, $x$ must satisfy the measurability condition. Q.E.D.

4.4. A SIMPLE SUFFICIENT CONDITION

We now present a simple sufficient condition under which measurability will be automatically satisfied. This condition is based on the possibility of small side payments and based on the idea of "scoring rules" (see Good (1952) and Winkler (1967)). Recall that the finest possible partitions is denoted by $\psi$. Let $\psi = \times_{i \in N} \psi_i$. Clearly any social choice function is measurable with respect to $\psi$ so that measurability with respect to $\psi$ is trivially satisfied.

Let $t_i \in [-\varepsilon, \varepsilon]$ denote the side payment to player $i$ and suppose that the total utility is $u_i(a, s) + t_i$. Let

$$p_i(\psi_i ; s_i) = \sum_{s_{-i} \in \psi_{-i}} p_i(s_{-i} ; s_i)$$

and suppose that for every $i \in N$, every $s_{-i} \in S_{-i}$, every $s_{1} \varepsilon S_{1}$, there exists $\psi_{-i} \varepsilon \psi_{-i}$ such that $p_i(\psi_{-i} ; s_{-i}) \neq p_i(\psi_{-i} ; s_{1})$. Construct a transfer rule $t_i : S_{1}$

$$\psi_{-1} \varepsilon \psi_{-1} \Rightarrow -\varepsilon, \varepsilon$$

such that
\[ t_1(s_1^*, \psi_{-1}^*) = -\alpha (1 - p_1(\psi_{-1}^*; s_1^*))^2 \sum_{\psi_{-1}^* \in \Psi_{-1}} \frac{p_1(\psi_{-1}^*; s_1^*)}{\psi_{-1}^* \in \Psi_{-1}} \]

where \( \alpha \) is a positive real number small enough to satisfy

\[ \varepsilon \leq t_1(s_1^*, \psi_{-1}^*) \leq 0 \text{ for all } s_1 \in S_1 \text{ and all } \psi_{-1}^* \in \Psi_{-1} \]

that for every \( s_1 \in S_1 \) and every \( s_{1'} \in S_1 \) (s,

\[ \sum_{\psi_{-1}^* \in \Psi_{-1}} (t_1(s_1^*, \psi_{-1}^*) - t_1(s_{1'}, \psi_{-1}^*)) p_1(\psi_{-1}^*; s_1^*) \leq 0 \]

consider "extended" social choice functions which map to \( A \times (-\varepsilon, \varepsilon) \). Let \( \Psi_{**} \) be the analogue of \( \Psi^* \) for "extended" social choice functions. It follows from the construction of transfer rule \( t_1 \) above that \( \Psi_{**} = \Psi^* \)
5. THE THEOREM

The main theorem asserts that self-selection and measurability are necessary and sufficient for implementation in iteratively undominated strategies. The necessity of self-selection and measurability were argued earlier and it only remains to establish sufficiency. The latter is true under two weak assumptions which are satisfied automatically if arbitrarily small fines may be levied. Before proceeding to details we provide a brief initial discussion premissed on the availability of small fines.

**Definition:** A social choice function \( x \) satisfies set self-selection for player \( i \) with respect to \( \psi \) if it is measurable with respect to \( \psi \) and

\[
U_i(x,s_i) \geq V_i(x,s_i,s_i') \quad \text{for all} \quad s_i \in S_i \quad \text{and} \quad s_i' \in S_i/(\psi_i(s_i))
\]

An social choice function \( x \) satisfies strict set self-selection for player \( i \) with respect to \( \psi \) if it satisfies set self-selection for player \( i \) with respect to \( \psi \) and the above inequalities strictly hold.

The strategy of proof is as follows. We first show that there exists a social choice function \( x \) which is strictly measurable with respect to \( \psi \) and is moreover exactly implementable in iteratively undominated strategies (this step is discussed below). This function is implementable using a mechanism in which players announce elements of their partition \( \psi_1^* \). Given this social choice function \( x \) we can virtually implement any social choice.
function \( x' \) which satisfies strict set self-selection for all players and is measurable with respect to \( x \), or equivalently measurable with respect to \( \Psi^* \) by mimicking the argument of Section 3 replacing announcements of signals by announcements of cells of the partition \( \Psi_i^* \): if a social choice function \( x \) satisfies the measurability condition, we will define a function \( x^+:\Psi^* \to A \) by
\[
x^+(\psi) = x(s) \text{ whenever } \psi = \gamma^*(s),
\]
where \( \gamma^*_i(s_i) \) is the element of \( \Psi^*_i \) that includes \( s_i \), and \( \gamma^*(s) = (\gamma^*_i(s_i))_{i \in N} \).

Let
\[
M_1 = \Psi_1^* \times \Psi_1^* \times \cdots \times \Psi_1^*,
\]
\[
g(m) = \varepsilon x^+(m^0) + (1 - \varepsilon) \frac{1}{K} \sum_{h=1}^{K} x^+(m^h).
\]

In addition, small fines are levied as before. The logic of the argument is now similar to the one presented earlier.

Note that for any social choice function \( x \) which satisfies the measurability condition and set self-selection for all players with respect to \( \Psi^* \), there exists a nearby social choice function \( x' \) which in addition satisfies strict set self-selection for all players with respect to \( \Psi^* \). This social choice function is
\[
nx + (1 - \eta)x
\]
Thus for virtual implementation, the distinction between weak and strict self-selection is unimportant.
How do we prove the existence of $x$? Recall the "dictatorial" function $f^1_i$ of Section 3. By an analogous argument, there exists a social choice function $x^1_i$ which satisfies strict set self-selection for player $i$ with respect to $\psi^1_i \times \psi^0_{-i}$. Hence the social choice function $\frac{1}{n} \sum_{i=1}^{n} x^1_i$ is strictly measurable with respect to $\psi^1_i$, and is implementable via (one round of) elimination of strictly dominated strategies.

Can we go a step further? By exactly the same logic there exists a social choice function $x^2_i$ which is measurable with respect to $\psi^2_i \times \psi^1_{-i}$ and satisfies strict set self-selection for player $i$. Then for small enough $\varepsilon$, the social choice function

$$\frac{1}{\varepsilon + \varepsilon} \left( \frac{\varepsilon}{n} \sum_{i=1}^{n} x^1_i + \frac{\varepsilon^2}{n} \sum_{i=1}^{n} x^2_i \right)$$

is implementable in iteratively strictly undominated strategies, and is strictly measurable with respect to $\psi^2_i$. We now need two rounds of iterative removal.

Proceeding in this way we inductively obtain the desired $x$. Now to details and formal proofs

**Assumption 1:** For every $i \in \mathbb{N}$, every $s_i \in S_i$, every $h \in (0, \ldots, L)$ and every $\psi_{-i} \in \psi_{-i}^h$, if $p_i(\psi_{-i}; s_i) > 0$, then there exist $a \in A$ and $a' \in A$ such that
\[ v = u_1(a,s) - u_1(a',s) p_1(s_{-1}; s_1) \]

where \( p_1(\psi_{-1}; s_1) = \sum_{s_{-1} \in \psi_{-1}} p_1(s_{-1}; s_1) \)

It is clear that this assumption is weak; it rules out indifference (in terms of conditional expected utility) across all lotteries. It is satisfied trivially if small transfers of private goods are permitted.

**Lemma 1:** For every \( i \in \mathbb{N} \) and \( h = 1, \ldots, L \), there exists a social choice function \( x^h_i \) which satisfies strict set self-selection for player \( i \) with respect to \( \psi^h_i \times \psi^{h-1}_{-i} \), where \( L \) is the positive integer introduced in Section 4 such that \( \psi^L = \psi^* \).

**Proof:** Fix \( i \in \mathbb{N} \) and \( h = 1, \ldots, L \) arbitrarily. Let \( \psi^{h-1}_i \) be the set of social choice functions which are measurable with respect to \( (S_i) \times \psi_{-i} \). Then \( s_i \) and \( s'_i \) induce different preference orderings over \( \psi^{h-1}_i \) if and only if \( \psi^h_i(s_i) \neq \psi^h_i(s'_i) \). By Assumption 1, neither of these preference orderings involves complete indifference. Hence, for all \( \psi_i \in \psi^h_i \) and \( \psi'_i \in \psi^h_i \), there exist social choice functions \( x \) and \( y \) which are measurable with respect to \( S_i \times \psi^{h-1}_{-i} \) such that...
\[ U_1(x, s_1) > U_1(y, s_1) \text{ if } \gamma_i^h(s_1) = \psi_i \]
\[ U_1(y, s_1) > U_1(x, s_1) \text{ if } \gamma_i^h(s_1) = \psi'_i \]

Let \( Z_i^h \) be a finite subset of \( \mathcal{Y}_i^{h-1} \) such that for all \( \psi_i \in \mathcal{Y}_i^h \) and \( \psi_i \not\in \mathcal{Y}_i^h/(\psi_i) \), there exist \( x \in Z_i^h \) and \( y \in Z_i^h \) with the above properties. Let \( J \)
\[ Z_i^h \]

For all \( \psi_i \not\in \mathcal{Y}_i^h \) and all \( j \in \{1, \ldots, J\} \) let \( x_j^i \) be social choice functions which satisfy \( \{x_j^i : j \in \{1, \ldots, J\}\} = Z_i^h \), and
\[ U_1(x_j^i, s_1) \geq U_1(x_{j+1}^i, s_1) \text{ for all } j = 1, 2, \ldots, J-1 \text{ if } \gamma_i^h(s_1) = \psi_i. \]

Consider the social choice function \( x_i^h \) defined by
\[ x_i^h(s) = \sum_{j=1}^{J} \alpha_j x_j^i(s) \text{ if } \gamma_i^h(s_1) = \psi_i, \]
where the \( \alpha_j \)'s are strictly positive, strictly decreasing in \( j \), and sum to one. Then \( x_i^h \) is measurable with respect to \( \mathcal{Y}_i^h \times \mathcal{Y}_{-i}^{h-1} \) and satisfies strict set self-selection for player \( i \)

Q.E.D.

**Lemma 2:** There exists a social choice function which is exactly implementable in iteratively undominated strategies and is strictly measurable with respect to \( \psi^* \)
PROOF: Define a social choice function \( x \) by

\[
x(s) = \alpha \sum_{l \in \mathbb{N}} \sum_{h=1}^{L} \varepsilon^{x_{1}^{h}}(s),
\]

where the social choice functions \( x_{1}^{h} \) are as in Lemma 1, and

\[
\alpha = \frac{1}{n \varepsilon^{2} + \cdots + \varepsilon^{L}}
\]

For small enough \( \varepsilon > 0 \), \( x \) is strictly measurable with respect to \( \psi^{*} \). We will show that \( x \) is exactly implementable.

We define a mechanism \( G = (M, g) \) by

\[
M_{i} = \psi_{i}^{*} \quad \text{for all } i \in \mathbb{N},
\]

and

\[
g = x^{*}
\]

We can choose a positive real number \( n > 0 \) such that for every \( i \in \mathbb{N} \), every \( s_{1} \in S_{1} \), every \( s_{1}' \in S_{1} \) and every \( h \in \{1, \ldots, L\} \), if \( s_{1}' \neq \psi_{1}^{h}(s_{1}) \), then

\[
U_{i}(x_{1}^{h}, s_{1}) - V_{i}(x_{1}^{h}, s_{1}', s_{1}') > n.
\]

For every \( x \in X \), define

\[
F_{i}(x) = \max_{(s, s')} \left\{ u_{i}(x(s'), s) - u_{i}(x(s'/s_{1}), s) \right\}.
\]

Choose \( \varepsilon \) small enough such that for every \( i \in \mathbb{N} \) and every \( h \in \{1, \ldots, L\} \),

...
Let \( \theta_1 \) and \( \theta_1' \) be strategies for agent 1 in \( G \).

Let \( P(h) \) be the statement: "if \( \theta \) is iteratively undominated, then for every \( \pi \in \Pi \) and every \( s_1 \in S_1 \), \( \theta_1(s_1) \) is a subset of \( \gamma_i^h(s_1) \)." Recall that \( \gamma_i^0(s_1) = S_1 \) for all \( s_1 \in S_1 \). Then \( P(0) \) holds trivially. We will show that for all \( h \in \mathbb{N} \), \( P(h) \Rightarrow P(h+1) \).

Suppose that \( P(h) \) and consider an iteratively undominated strategy \( \theta_1 \) for player 1. Fix \( s_1 \in S_1 \) arbitrarily. Suppose \( \theta_1(s_1) \) is a subset of \( \gamma_i^h(s_1) \) and \( \theta_1'(s_1) \) is a subset of \( \gamma_i^{h-1}(s_1) \) and also a subset of \( \gamma_i^h(s_1') \) for some \( s_1' \in S_1/(\gamma_i^h(s_1)) \). Then

\[
\begin{align*}
& v_1(G, \theta, s_1) - v_1(G, \theta, s_1') \\
& \geq \sum_{j \in N} \sum_{k=h+1}^{L} \epsilon_k F_1(x_j^k) \\
& \geq \sum_{j \in N} \sum_{k=h+1}^{L} \epsilon_k F_1(x_j^k) \\
& \geq \varepsilon \alpha(U_1(x_1^h, s_1) - V_1(x_1^h, s_1, s_1') - \sum_{j \in N} \sum_{k=h+1}^{L} \epsilon_k F_1(x_j^k) \\
& \Rightarrow 0.
\end{align*}
\]

means that if \( \theta_1 \) is iteratively undominated, \( \theta_1(s_1) \) is a subset of \( \gamma_i^h(s_1) \), and therefore, \( P(h) \) must hold.
Hence $g = \gamma^*$. Finally recall that

$$g(\gamma^*(s)) = x(s) \text{ for all } s \in S.$$  

Q.E.D.

ASSUMPTION 2: For every $i \in \mathbb{N}$ and every $\psi \in \Psi^*_\ast$, there exist $a(i, \psi) \in A$ and $a(i, \psi) \in A$ such that for every $\Psi \in \Psi$, if $p_i(s_{-i}; s_i) > 0$, then

$$u_i(a(i, \psi), s) - u_i(a(i, \psi), s) > 0$$

and

$$u_j(a(i, \psi), s) - u_j(a(i, \psi), s) \geq 0 \text{ for all } j \in \mathbb{N}/\{i\}$$

Assumption 2 corresponds to the single assumption used in our earlier paper on complete information. Like Assumption 1, Assumption 2 is trivially satisfied if strictly positive (though possibly, arbitrarily small) transfers of private goods are possible. It requires that for every player $i$ and every signal profile, there exist a pair of lotteries which are strictly ranked for player $i$ and for which other players have the (weakly) opposite ranking. In addition these pair of lotteries must be chosen measurably with respect to $\Psi^*_\ast$. 
THEOREM: A social choice function is virtually implementable in iteratively undominated strategies if and only if it satisfies self-selection and is measurable with respect to $\psi^*$.

PROOF: As noted above self-selection is obviously necessary. By Proposition 1, so is measurability. Hence, it only remains to establish the 'if' part of the theorem.

By Lemma 2, there exists a social choice function $x$ which is strictly measurable with respect to $\psi^*$ and exactly implementable in iteratively undominated strategies. Let $G = (M, g)$ be the implementing game form constructed in the proof of Lemma 2, where $M_i = \psi_i^*$ for all $i \in N$, $g = x^*$, and $\gamma^*$ is the unique iteratively undominated strategy profile of $G$. Since $\gamma^*$ is a strict Bayesian Nash equilibrium in $G$, $x$ satisfies strict set self-selection for all players with respect to $\psi^*$. Let

$$\gamma(s) = (1 - \alpha)x(s) + \alpha x(s).$$

Since $x$ satisfies self-selection and $x$ satisfies strict set self-selection all players with respect to $\psi^*$, $\gamma^*$ satisfies strict set self-selection all players with respect to $\psi^*$ also, for any $\alpha \in (0, 1]$. We will assume without loss of generality that $x^*$ satisfies strict set self-selection for players with respect to $\psi^*$ since there exist arbitrarily close social choice functions which do.
The game form $G = (M, g)$ we construct here is similar to the one for complete information. Specifically,

$$M_1 = M_1^0 \times M_1^1 \times \cdots \times M_1^K = \psi_1^* \times \cdots \times \psi_1^*$$

for an integer $K$ to be defined.

Define the function $\varepsilon : N \times M \to A$ as follows:

$$\varepsilon(i, m) = a(i, m^0) \quad \text{if there exists } k \in \{1, \ldots, K\} \text{ such that } m_k^0 \neq m_i^0 \text{, and } m^h = m_i^0 \text{ for all } h \neq k, k-1$$

$$\varepsilon(i, m) = a(i, m^0) \quad \text{otherwise,}$$

where $a$ and $a'$ are as in Assumption 2.

The outcome function $g : M \to A$ is

$$g(m) = \varepsilon x^+(m^0) + \varepsilon \sum_{i \in N} \varepsilon(i, m) + (1 - \varepsilon - \varepsilon^2) \sum_{h=1}^{K} x^+(m^h)$$

Define $\sigma_i^* = \psi_i^* \ldots, \psi_i^*$, and let $\sigma^* = (\sigma_i^*)_{i \in N}$. We show that for small enough $\varepsilon$, $\sigma^*$ is the unique iteratively undominated strategy profile in the game form $G$.

We first argue that if $\sigma$ is iteratively undominated in $G = (M, g)$, then $\sigma^0 = \psi^*$. For every $i \in N$, every $\sigma$ and every $\sigma'$, if $\sigma_i^h = \sigma_i^h$ for all $h \neq 1, \ldots, K$, then

$$v_1(G, \sigma/\sigma_i^0, s_i) \leq v_1(G, \sigma, s_i)$$
For the mechanism \( G \), let \( (Q^h)_{1 \in \mathbb{N}} \), \( h = 1,2,\ldots, \) be a sequence of sets of iteratively undominated strategies as defined in Section 3, and let \( (Q^h_{1})_{1 \in \mathbb{N}} \), \( h = 1,2,\ldots, \) be the corresponding sequence for \( G \). It follows from definitions that there exists a positive number \( n > 0 \) such that for every \( 1 \in \mathbb{N} \), \( h \in \{0,1\} \), and \( \theta_1 \vDash Q^h_{1} \), if \( \theta_1 \) is dominated with respect to \( Q^h_{1} \), then there exist \( s_1 \vDash S_1 \) and \( \theta_1' \vDash Q^h_{1} \) such that for all \( \theta_1 \vDash Q^h_{1} \),

\[
\varepsilon(\nu_1(G,\theta_1',s_1) - \nu_1(G,\theta_1,s_1)) \geq \frac{2\varepsilon^2}{n} E_1
\]

For every \( 1 \in \mathbb{N} \), let

\[
E_1 = \max \{ \sum_{s \vDash S, m \in M} u_1(\xi(j,m),s) : j \in \mathbb{N} \}
\]

Choose \( \varepsilon > 0 \) such that

\[
\varepsilon \leq \frac{2\varepsilon^2}{n} E_1 \quad \text{for all } 1 \in \mathbb{N},
\]

and assume below that \( 0 < \varepsilon \leq \varepsilon \). Then

\[
\varepsilon(\nu_1(G,\theta_1',s_1) - \nu_1(G,\theta_1,s_1)) \geq \frac{2\varepsilon^2}{n} E_1
\]
\[ \varepsilon(v_1(G, \sigma^0, s_1, q_1) - v_1(G, \sigma^0, s_1) - q \]

It follows that if \( \sigma \) is an element of \( Q^1 \), then \( \sigma^0 \) must be an element of \( Q^1 \)
and recursively, for every \( h \in \{2, \ldots \} \), if \( \sigma \) is an element of \( Q^h \), then \( \sigma^0 \) must
be an element of \( Q^h \). Consequently, if \( \sigma \) is iteratively undominated in \( G \)
then, \( \sigma^0 = \gamma^* \)

For every \( i \in \mathbb{N} \) and every \( s \in S \), define

\[
B_i(s) = u_i(a(i, \gamma^*(s)), s) - u_i(a(i, \gamma^*(s)), s),
\]

\[
D_i(s) = \max_{s' \in S} [u_i(x(s), s) - u_i(x(s/s_1'), s) + u_i(x(s'), s)
- u_i(x(s'/s_1), s)]
\]

By Assumption 2, for every \( i \in \mathbb{N} \) and every \( s \in S \)

\[ B_i(s) > 0 \text{ if } p_i(s_{-i}, s_i) > 0. \]

Hence, there exists a positive integer \( K \) such that for every \( i \in \mathbb{N} \) and every \( s \in S \),

\[ K \frac{\varepsilon}{n} B_i(s) > (1 - \varepsilon - \varepsilon^2)D_i(s) \text{ if } p_i(s_{-i}, s_i) > 0. \]

Let \( P(h) \) be the statement: "if \( \sigma \) is iteratively undominated in \( G \), then
for every \( i \in \mathbb{N} \)

\[ q_{\sigma} = \gamma^* \text{ for all } q_{\varepsilon}(0, \ldots, h). \]

We have established \( P(0) \). We now show that for all \( h = 1, \ldots, K \), \( P(h-1) \supset P(h) \).
Suppose $P(h-1)$ and consider an iteratively undominated strategy $\gamma_1$ for player 1. Then, by $P(h-1)$,
$$\gamma_1^* \text{ for all } q \in \{0, \ldots, h-1\}$$
we need to show that $\gamma_1^* = \gamma_1^*$ also. Suppose not, and let $\gamma_1^*(s_1) = \gamma_1^*(s_1')$ for some $s_1 \in S_1$. Let $\gamma_1^*$ be the strategy for player 1 such that
$$\gamma_1^* \text{ for all } q \in h, \text{ and } \gamma_1^*(s_1) = \gamma_1^*(s_1)$$
Consider any iteratively undominated strategy profile $\sigma_{-1}$ for other players.

Then $\sigma_j = \gamma_j^*$ for all $q \in \{0, \ldots, h-1\}$ and all $j \in N\setminus\{1\}$. Let
$$s_{-1} \in S_{-1}, \quad \sigma_{-1}^h(s_{-1}) \neq \gamma_{-1}^*(s_{-1})$$
Then, by $(**)$
$$v_1(G, \sigma, \sigma_1, s_1) = v_1(G, \sigma, s_1)$$
$$= \frac{1}{K} \sum_{s_{-1} \in S_{-1}} \left( 1 - \varepsilon - \varepsilon^2 \right) \left( u_1(x(s), s) - u_1(x(s/s_1^1), s) \right) p_1(s_{-1}; s_1)$$
$$+ \frac{1}{K} \sum_{s_{-1} \in S_{-1}} \left( 1 - \varepsilon - \varepsilon^2 \right) \left( u_1(x^+(s), \gamma_1^*(s_1)) , s \right)$$
$$- u_1(x^+(s), s)) + \frac{2}{n} B_1(s)p_1(s_{-1}; s_1)$$
$$\frac{1}{K} \left( 1 - \varepsilon - \varepsilon^2 \right) (U_1(x, s_1) - v_1(x, s_1, s_1'))$$
$$+ \frac{1}{K} \sum_{s_{-1} \in S_{-1}} (K\frac{2}{n} B_1(s) - 1 - \varepsilon - \varepsilon^2) D_1(s)p_1(s_{-1}; s_1)$$
The extra subtlety here, which does not appear in the complete information case, is that some types of other players (this is the set $S_{-1}^*$) may misrepresent $m_1^h$ while others do not.

This inequality implies that $n_1$ is not iteratively undominated in $G$, a contradiction. Hence, $n_1^h(s_1) = s_1^*$, $n_1^h = s_1^*$, and the unique iteratively undominated strategy profile is $s^* = s^*_{-1}, ..., s^*$.

Since
\[
g(s^*) = 1 - \varepsilon - \varepsilon^2 x^+(s^*) + \varepsilon x^+(s^*)
\]
\[
+ \frac{\varepsilon^2}{n} \sum_{i \in N} a(i, s^*(s)),
\]
and $\varepsilon$ can be taken to be arbitrarily small, the proof is complete.

Q.E.D.
6. COMPLETE INFORMATION AND THE TWO-PLAYER CASE

Incomplete information framework is very general and in particular incorporates the special case of complete information. Furthermore our result is independent of the number of players and in particular covers the two-player case, which traditionally has been treated separately (or not at all) with equilibrium-based implementation literature. We spell out here how our theorem specializes to the complete information case with two or more players.

In general, some signals \( s = (s_1, \ldots, s_n) \in S \) may occur with zero probability. This is, for instance, the case in the complete information setting where all players receive the same signal that is all probability mass lies on the "diagonal" of \( S \) (i.e., \( s \in S \) such that \( s_1 = s_2 \) for all \( i \in \{1, \ldots, n\} \)).

The self-selection condition is expressed in terms of social choice functions which are defined for all \( s \in S \). Of course, from the point of view of final realized equilibrium outcomes, any pair of social choice functions which agree on \( S \), the support of \( S \), are equivalent. Starting from a social choice function defined on \( S \) we would require that the self-selection condition be satisfied for some extension of the function to \( S \), the cartesian product of the individual signal space \( S_1 \).

In our paper on complete information, Abreu and Matsushima (1992), it was convenient to define the social choice function on \( S \), i.e., the diagonal
of $S$. With three or more players such a social choice function may be extended to $S$ in a manner in which the self-selection condition is automatically satisfied; the extension simply ignores a single player deviation from an otherwise unanimous announcement.

In terms of the notation of this paper the complete information case may be described as follows:

$$p_i(s_{-i}; s_i) = \begin{cases} 1 & \text{if } s_j = s_i \text{ for all } j \in N/(i), \\ 0 & \text{otherwise,} \end{cases}$$

$$S = \{(s_1, \ldots, s_n) \in S : s_1 = s_i \text{ for all } i \in N\}.$$

For any $s_i \in S_i$, let $s_i^e$ denote $t \in S$ for which $t_j = s_i$ for all $j \in N$. Each $s_i \in S_i$ corresponds to a preference profile, one for each player. Each $S_i$ may be partitioned into subsets $\phi_i(s_i)$ where each element of $\phi_i(s_i)$ yields the same preferences over lotteries for player $i$. Furthermore, for all $s_i \in S_1$, $(s_i) = \bigcap_{j \in N} \phi_j(s_i)$.

We first argue that measurability is trivial in complete information environments. In fact, this follows from our earlier paper (at least when there are three or more players) which shows that any social choice function is virtually implementable in the complete information case. Since measurability is a necessary condition, it follows that with complete information any (three or more player) social choice function is measurable. We sketch a direct argument below (which also covers the two-player case).
Specialized to the complete information case Assumption 1 amounts to ruling out player types who are indifferent over all lotteries. As we show in AM this implies that there exists a function \( f_1: S_1 \rightarrow A \) such that for every \( s_1 \in S_1 \),

\[
f_1(s_1) = f_1(s'_1) \quad \text{for all } s'_1 \in \Phi(s_1)
\]

\[
u_1(f_1(s_1), s_1) > u_1(f_1(s'_1), s_1) \quad \text{for all } s'_1 \not\in \Phi(s_1).
\]

By definition

\[\psi^0_1 = \{s_1\}\]

By considering constant social choice functions \( x^1_{t_1}: S \rightarrow A \) where

\[
x^1_{t_1}(z) = f_1(t_1) \quad \text{for all } z \in S,
\]

the reader may check that \( R^1(\psi^0_{-1}) = \{\phi_1(s_1) : s_1 \in S_1\} \) (for any \( w_1 \not\in \Phi(s_1) \)). Consider the pair of constant social choice functions \( x^1_{s_1} \) and \( x^1_{w_1} \). Hence

\[\psi^1_1 = \{\phi_1(s_1) : s_1 \in S_1\}.
\]

Finally the second round of the iterative procedure yields

\[\psi^2_1 = \psi^1_1 = \{s_1) : s_1 \in S_1\}
\]

the finest possible partition. Hence \( \psi^*_1 = \psi^2_1 \), and the iterative procedure terminates in the second round. Let \( B_1(\phi_1(s_1)) \) and \( b_1(\phi_1(s_1)) \) satisfy
existence of such lotteries is guaranteed by Assumption 1 (non-indifference over all lotteries). In distinguishing between any pair of signals \( w_i \) and \( s_i \) we may now use social choice functions which are measurable with respect to \( (S_i \times S_i^{-1}) \). Define \( g_1 : S \to A \) by

\[
g_1(s) = \begin{cases} b_1(\phi_1(s)) & \text{if } \phi_j(s_j) = \phi_j(s_i) \text{ for all } j \in N, \\ b_i(\phi_1(s)) & \text{otherwise.} \end{cases}
\]

Let \( x^2_{t_1} : S_i \to A \) be defined by

\[
x^2_{t_1}(s) = g_1(t_1,s_{-1}) \text{ for all } s \in S.
\]

to distinguish between \( w_i \) and \( s_i \) where \( w_i \sim s_i \) and \( w_i \nsim \phi_1(s_i) \), consider social choice functions \( x^2_{s_1} \) and \( x^2_{w_1} : U_1(x^2_{s_1},s_1) = U_1(b_1(s_1),s_1) \)

and conversely when the signal is \( w_i \).

Let us turn to the self-selection condition. As noted above this condition is satisfied trivially when there are three or more players. Let \( x : S \to A \) be a social choice function as in the earlier paper on complete information. It may be extended to yield an equivalent (on the support \( S \) of \( S \)) social choice function \( x : S \to A \) which satisfies self-selection as follows:

\[u_1(b_1(s_1),s_1) = U_1(x^2_{w_1},s_1)\]
\[ x(s) = x(s_j \cdot e) \quad \text{if there exists a subset } J \text{ of } N \text{ such that } J; \geq n-1, j \in J \text{ and } s_1 = s_j \text{ for all } i \in J, \]

\[ x(s) = b \quad \text{otherwise, where } b \text{ is an arbitrary element of } A. \]

This extension is not well defined if \( n \geq 3 \). Indeed in the two-player case, self-selection is a non-trivial requirement, and is equivalent to a condition termed the **intersection** property.

**DEFINITION:** A two-person social choice function \( x : S \rightarrow A \) satisfies the **intersection** property if for every \( s_1 \in S_1 \) and \( s_2 \in S_2 \) there exists \( b \in A \) such that

\[ u_1(x(s_2 \cdot e), s_2 \cdot e) \geq u_1(b, s_2 \cdot e), \]

and

\[ u_2(x(s_1 \cdot e), s_1 \cdot e) \geq u_2(b, s_1 \cdot e). \]

When this property is satisfied the set of lotteries which is (weakly) worse than \( x(s_j \cdot e) \) for player 1 with preferences \( \gamma_1(s_j) \) has a non-empty intersection with the corresponding set for player 1. For a social choice function \( x \) which satisfies the intersection property and each \( s_i \in S_i \) and
let \( b(s_1, s_2) \) satisfy the inequalities of the definition. Define an equivalent extended social choice function \( x: S \to A \) by

\[
x(s_1, s_2) = x(s_1) \quad \text{if} \ s_1 = s_2,
\]

\[
x(s_1, s_2) = b(s_1, s_2) \quad \text{if} \ s_1 \neq s_2.
\]

\( x \) satisfies self-selection. It is also clear that \( x \) must satisfy the intersection property if some extension of it satisfies self-selection. We remark that the existence of a "holocaust" outcome is a crude sufficient condition for the intersection property to hold.

It follows immediately from our theorem and the preceding discussion in the complete information case, any two-person social choice function which satisfies the intersection property is virtually implementable in iteratively undominated strategies.

Abreu and Sen (1989) showed that the intersection property is necessary (and sufficient for virtual implementation in Nash equilibrium). We note that the intersection property is much weaker than the complex necessary and sufficient conditions for exact two-person implementation in Nash equilibrium provided by Dutta and Sen (1988) and Moore and Repullo (1988).
CONCLUSION

A large literature following the Gibbard-Satterthwaite theorem has sought to characterize implementable social choice functions for a variety of concepts weaker than dominant strategy implementation. Under general informational assumptions, and in the context of virtual implementation and iterative elimination of strictly dominated strategies this paper provides a characterization which is essentially complete. We present two necessary conditions, self-selection and measurability, which (under weak domain restrictions) are also sufficient. The former is obvious and well-known, and the latter is introduced in the present paper. Measurability is frequently automatically satisfied. Modulo measurability, our result is therefore as permissive as one could possibly expect, and furthermore obtained for a solution concept which is even weaker than rationalizability.

We are aware of no characterizations involving non-equilibrium solution concepts in general Bayesian environments. Earlier results for implementation in Bayesian Nash equilibrium are limited by the stringency of the necessary condition (Bayesian monotonicity), the restriction to three or more players, the (undesirable) equilibrium basis of the solution concept above all by the unsatisfactory nature of the implementing mechanisms.
1. See Section 5 in Abreu and Matsushima (1994) for a discussion of this assumption.

2. See however Matsushima (1993c).

3. See Matsushima (1994a), who argues that Bayesian monotonicity is trivial when side payment is permitted.

4. Palfrey and Srivastava (1989a) dispense with this condition for the special case of private values by considering implementation in undominated Bayesian Nash equilibrium. Their mechanisms are, however, suspect in that they involve the elimination of dominated strategies in unbounded mechanisms. See Jackson (1989b). In such mechanisms a strategy x may be dominated even though there exists no undominated strategy y which dominates x. Such perversities appear to play a key role in obtaining the desired results.

5. Most of the mechanisms that have been used in the Nash and Bayesian Nash literature are not regular. We take this to be additional evidence of the unsatisfactory nature of these mechanisms.
Suppose that $x$ is strictly measurable with respect to some partition $\psi$ and $y$ is measurable with respect to some coarser partition $\psi'$. Then $(1-\lambda)x + \lambda y$ need not be strictly measurable with respect to $\psi$ for arbitrary $\lambda$, but will be for small enough $\lambda$.

Since measurability is trivially satisfied in complete information environments whereas not all social choice functions are monotonic in such environments, it is clear that the measurability condition does not imply Bayesian monotonicity.
REFERENCES


