

Serrano, Roberto; Vohra, Rajiv

**Working Paper**

## Multiplicity of mixed equilibria in mechanisms: A unified approach to exact and approximate implementation

Working Paper, No. 2009-11

**Provided in Cooperation with:**

Department of Economics, Brown University

*Suggested Citation:* Serrano, Roberto; Vohra, Rajiv (2009) : Multiplicity of mixed equilibria in mechanisms: A unified approach to exact and approximate implementation, Working Paper, No. 2009-11, Brown University, Department of Economics, Providence, RI

This Version is available at:

<https://hdl.handle.net/10419/62639>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Multiplicity of Mixed Equilibria in Mechanisms: a Unified Approach to Exact and Approximate Implementation\*

Roberto Serrano<sup>†</sup> and Rajiv Vohra<sup>‡</sup>

Department of Economics  
Brown University  
Providence, RI 02912, U.S.A.

This version: September 2009

**Abstract.** We characterize full implementation of social choice sets in mixed-strategy Bayesian equilibrium. Our results concern both exact and virtual mixed implementation. For exact implementation, we identify a strengthening of Bayesian monotonicity, which we refer to as mixed Bayesian monotonicity. It is shown that, in economic environments with at least three agents, mixed Bayesian implementation is equivalent to mixed Bayesian monotonicity, incentive compatibility and closure. For implementing a social choice function, the case of two-agents is also covered by these conditions and mixed Bayesian monotonicity reduces to Bayesian monotonicity. Following parallel steps, mixed virtual implementation is shown to be equivalent to mixed virtual monotonicity, incentive compatibility and closure. The key condition, mixed virtual monotonicity, is argued to be very weak. In particular, it is weaker than Abreu-Matsushima's measurability, thereby implying that: (1) virtual implementation in mixed Bayesian equilibrium is more permissive than virtual implementation in iteratively undominated strategies, and (2) non-regular mechanisms are essential for the implementation of rules in that gap. JEL Classification: C72, D78, D82.

Keywords: Exact implementation, approximate implementation, incomplete information, incentive compatibility, monotonicity.

---

\*We thank Georgy Artemov, Takashi Kunimoto and Stephen Morris for helpful comments. Some of the results reported here were first presented at the 2004 North American Summer Meeting of the Econometric Society at Brown. We acknowledge support from NSF grant SES-0133113. Serrano also thanks Deutsche Bank for research support, and the Institute for Advanced Study in Princeton and CEMFI in Madrid for their hospitality.

<sup>†</sup>Department of Economics, Brown University and IMDEA-Social Sciences, roberto.serrano@brown.edu; [www.econ.brown.edu/faculty/serrano](http://www.econ.brown.edu/faculty/serrano)

<sup>‡</sup>Office of the Dean of the Faculty, Brown University, rajiv\_vohra@brown.edu; [www.econ.brown.edu/~rvohra](http://www.econ.brown.edu/~rvohra)

# 1 Introduction

The literature on implementation with incomplete information has often left out the consideration of mixed-strategy equilibria. This is particularly problematic for a research program that attempts to address the problem of multiplicity of equilibria in economic institutions. Thus, the current work attempts to close an important gap in the implementation literature. It provides characterizations of the social choice rules that can be decentralized by means of all the mixed-strategy equilibria of a mechanism. In doing so, the paper takes a unified approach to exact and approximate implementation, and considers general social choice sets in economic environments with incomplete information – outside of economic environments, tight characterizations are generally not available, even for pure-strategy equilibria.

Most of the papers on implementation that use equilibrium concepts have confined their scope to the case of pure strategies.<sup>1</sup> This includes the contributions to virtual or approximate nash implementation (Abreu and Sen (1991), Matsushima (1988)), the results for exact Bayesian implementation (Postlewaite and Schmeidler (1986), Palfrey and Srivastava (1987, 1989), Mookherjee and Reichelstein (1990), Jackson (1991)) and those on virtual Bayesian implementation (Serrano and Vohra (2001, 2005)). This restriction is undesirable and it makes comparisons with other results more difficult. In particular, we are thinking of the connections with implementation using undominated or iteratively undominated strategies (Jackson (1992), Abreu and Matsushima (1992a, b)), whose results have implications for mixed-strategy equilibrium implementation, as well as for the mechanisms employed.

Our unified approach to exact and approximate implementation is facilitated by the use of random mechanisms. That is, even when we deal with exact implementation, we consider allocation rules that map information states into probability distributions over alternatives. This is especially natural if one considers incomplete information environments, as we do, since agents are already involved in problems of decision making under uncertainty, given their asymmetric information. In addition, we shall relax the assumption of finite type spaces, made by most of the papers mentioned above.<sup>2</sup>

In Theorem 1, we show that incentive compatibility, mixed Bayesian

---

<sup>1</sup>Two salient exceptions are Duggan (1997) and Maskin (1999). Duggan provides a sufficiency theorem for virtual Bayesian implementation, while Maskin restricts attention to complete information environments.

<sup>2</sup>Duggan (1997) is again the exception here.

monotonicity and closure are necessary and sufficient for a social choice set to be Bayesian implementable in mixed strategies when there are at least three agents. With respect to Jackson’s (1991) result characterizing Bayesian implementation in pure strategies for the same problems, the only difference stems from the two monotonicity conditions. While his Bayesian monotonicity considers only pure deceptions, mixed Bayesian monotonicity is formulated for all mixed deceptions. Suppose  $F$  is a social choice set, and consider a function  $f \in F$ . Let  $\alpha$  be a mixed deception, i.e., a (possibly random) play of the direct mechanism for  $f$  that is different from truth-telling. Mixed Bayesian monotonicity says that if a mixed deception  $\alpha$  undermines our goals, i.e.,  $f \circ \alpha \notin F$ , then there must exist an agent type  $t_i$  and another function  $y$  that exhibit a preference reversal (while he prefers  $f$  over  $y$ , he prefers  $y \circ \alpha$  over  $f \circ \alpha$ ). Importantly, we show that mixed Bayesian monotonicity reduces to Bayesian monotonicity if one considers social choice functions instead of multivalued sets. The reason is that any strictly mixed deception exhibiting such a preference reversal can be “decomposed” into the pure deceptions in its support exhibiting a similar reversal.<sup>3</sup> However, when one considers multivalued sets, mixed Bayesian monotonicity is more restrictive: it only reduces to Bayesian monotonicity when the set satisfies a convex range property over the set of pure deceptions.

Theorem 2 studies social choice functions and also covers the two-agent case. With functions, closure is a trivial requirement and mixed Bayesian monotonicity reduces to Bayesian monotonicity. Therefore, incentive compatibility and Bayesian monotonicity provide the full characterization for this case.

For approximate implementation with incomplete information, social choice sets have not been treated so far. Serrano and Vohra (2005) obtained a characterization of social choice functions and implementability in pure-strategy equilibria. Our Theorems 3 and 4 are exact parallels of Theorems 1 and 2, respectively. The only change in the conditions is the replacement of mixed Bayesian monotonicity with mixed virtual monotonicity, which is the extension to mixed deceptions of Serrano and Vohra’s (2005) virtual monotonicity condition. Thus, mixed virtual monotonicity requires the kind of preference reversal described above, but involving an incentive compatible function  $x$

---

<sup>3</sup>A different reason to see that approaches based on pure versus mixed equilibria should make no difference is found in robust implementation, where results are sought to be robust over multiple type spaces (see, e.g., Bergemann and Morris (2009) and Artemov, Kunimoto and Serrano (2007)).

(not necessarily in  $F$ ) and another function  $y$  for each mixed deception  $\alpha$ . The same comments made for mixed Bayesian monotonicity of sets versus functions apply to mixed virtual monotonicity, which can also be shown to be an extremely permissive condition, as discussed in Subsection 5.2.

Abrue and Matsushima (1992b) characterize the functions that are virtually implementable in iteratively undominated strategies in terms of incentive compatibility and a measurability condition. They also argue how the same conditions characterize virtual implementation in mixed-strategy equilibria if one relies on small monetary punishments out of equilibrium and uses regular mechanisms.<sup>4</sup> An important example due to Duggan (1997) exhibits a function in a finite environment that is not measurable in the AM sense, but that can be exactly implemented in the standard sense, i.e., in pure-strategy Bayesian equilibria, by means of its direct mechanism. We further analyze this example and, in light of our results, we can conclude that non-regular mechanisms are essential to the approximate Bayesian implementation of functions that are not AM-measurable.

This is the plan of the paper. Section 2 goes over the model and its basic definitions. Section 3 is concerned with exact implementation, and the parallel results on approximate implementation are found in Section 4. Section 5 discusses the connections between our work, Maskin's approach to mixed-strategy Nash implementation and AM-measurability.

## 2 The Model and Definitions

We shall consider implementation in the context of a general environment with asymmetric information. Let  $N = \{1, \dots, n\}$  be a finite set of agents. Let  $T_i$  denote the set of agent  $i$ 's types; these will be arbitrary sets, including of uncountably infinite cardinality. The interpretation is that  $t_i \in T_i$  describes the *private information* possessed by agent  $i$ . We refer to a profile of types  $t = (t_1, \dots, t_n)$  as a state. Let  $T = \prod_{i \in N} T_i$  be the set of states. We will use the notation  $t_{-i}$  to denote  $(t_j)_{j \neq i}$ . Similarly  $T_{-i} = \prod_{j \neq i} T_j$ . The set of types  $T_i$  is equipped with the  $\sigma$ -algebra  $\mathcal{T}_i$ . Let  $\mathcal{T}$  denote the product  $\sigma$ -algebra over  $T$ .<sup>5</sup>

---

<sup>4</sup>In regular mechanisms best responses always exist; integer games, open strategy sets, and devices alike are hence ruled out.

<sup>5</sup>If  $T$  is countable, the natural  $\sigma$ -algebra is discrete, i.e., the one containing all subsets of  $T$ . In this case, strategies are automatically well behaved. For more general spaces,

Let  $A$  denote the set of social alternatives, which is assumed to be independent of the information state. Let  $\mathcal{A}$  be a  $\sigma$ -algebra on  $A$  containing all singleton sets, and  $\Delta$  denote the set of probability measures on  $(A, \mathcal{A})$ .

The Bernoulli utility of agent  $i$  for pure alternative  $a$  in state  $t$  is  $u_i(a, t)$ . Abusing notation slightly, given  $B \in \mathcal{A}$ ,  $u_i(B, t)$  will refer to agent  $i$ 's expected utility evaluation of lottery  $B$  in state  $t$ . We shall assume that  $u_i$  is  $\mathcal{A} \times \mathcal{T}$ -measurable.

Each agent's prior belief about the state is given by a probability measure  $\mu_i$  on  $(T, \mathcal{T})$ . These beliefs agree on sets of measure 0: for all  $i, j \in N$  and for all  $S \in \mathcal{T}$ ,  $\mu_i(S) = 0$  if and only if  $\mu_j(S) = 0$ . Let  $T^* \subseteq T$  be the smallest set of states with full measure, i.e.,  $\mu_j(S) = 0$  for every  $S$  for which  $S \cap T^* = \emptyset$ .

We can now define an *environment* as  $\mathcal{E} = \{(A, \mathcal{A}), (u_i, (T_i, \mathcal{T}_i), \mu_i)_{i \in N}\}$ .

A *social choice function* (SCF) is a mapping  $f : T \mapsto \Delta$ . The notation  $f(B, t)$  refers to the likelihood assigned to  $B \in \mathcal{A}$  by the measure  $f(t)$ . We assume that  $f(B, t)$  is  $\mathcal{T}$ -measurable as a function of  $t$ . Denote the set of SCFs by  $\mathbf{F}$ , equipped with  $\sigma$ -algebra  $\mathcal{F}$ .

A *social choice set* (SCS)  $F$  is a subset of  $\mathbf{F}$ .

Consider the following metric on SCFs:

$$d(f, h) = \sup\{|f(A' | t) - h(A' | t)| \mid t \in T^*, A' \in \mathcal{A}\}.$$

For  $\epsilon \geq 0$ , we shall say that two SCSs  $F$  and  $H$  are  $\epsilon$ -approximate ( $F \approx_\epsilon H$ ) if there exists a bijection  $\pi$  such that for every  $f \in F$  and every  $h \in H$   $h = \pi(f)$  satisfying that  $d(f, \pi(f)) \leq \epsilon$  for every  $f \in F$ .

We shall say that two SCSs  $F$  and  $H$  are *equivalent* ( $F \approx H$ ) if they are 0-approximate. This means that the two sets “coincide” for every  $t \in T^*$ .

Note that since all measures  $\mu_i$  are absolutely continuous, i.e., agree on the identification of sets of measure 0, these concepts are well defined. Also, for any two SCFs  $f$  and  $h$ , we shall assume that the set  $S \subseteq T$  such that if  $t \in S$   $f(t) \neq h(t)$  is  $\mathcal{T}$ -measurable: a host of different sufficient conditions can be used to guarantee this; for example, those identified in Duggan's (1997) Proposition 1.

A *mechanism*  $G = ((M_i, \mathcal{M}_i)_{i \in N}, g)$  consists of a message space  $M_i$  for agent  $i$ , equipped with a  $\sigma$ -algebra  $\mathcal{M}_i$ , where we denote by  $M$  the Cartesian product of the  $M_i$ s and by  $\mathcal{M}$  the product  $\sigma$ -algebra, and an outcome function

---

the restriction to measurable strategies is called for. While it may sometimes be judged as counterintuitive, it is a technical requirement for agents to be able to calculate best responses. The reader is referred to Duggan's (1997) introduction for a clear explanation.

$g : M \mapsto \Delta$ . Thus,  $g(m)$  is an SCF, and we denote by  $g(m(A', t))$  the likelihood assigned to  $A' \in \mathcal{A}$  by the measure  $g(m(t))$ . We shall assume that  $g(m)$  is  $\mathcal{M}$ -measurable as a function of  $m$ .

Denote by  $\Delta(M_i)$  the set of measures (mixed strategies) over  $(M_i, \mathcal{M}_i)$ . In general, given a measurable mixed strategy profile  $\tilde{m} \in \prod_{i \in N} \Delta(M_i)$ , we shall assume that  $g(\tilde{m}(t))$  is  $\mathcal{M} \times \mathcal{T}$ -measurable. Therefore, for any  $m_i \in M_i$ ,  $g(m_i, \tilde{m}_{-i}(t))$  is  $\mathcal{M}_{-i} \times \mathcal{T}_{-i}$ -measurable.

For each  $i \in N$  and  $t_i \in T_i$ , the conditional likelihood of a set  $S_{-i} \in \mathcal{T}_{-i}$ , given  $t_i$  is denoted  $\mu_i(S_{-i} | t_i)$ .

The (interim/conditional) expected utility of agent  $i$  of type  $t_i$  corresponding to an SCF  $f$ , whenever it exists, is defined as:

$$U_i(f|t_i) \equiv \int_{T_{-i}} \int_A u_i(f(a, t), t) \mu_i(S_{-i}|t_i) da dS_{-i}.$$

A (mixed) *Bayesian equilibrium* of  $G$  is a profile of strategies  $(\tilde{m}_i)_{i \in N}$  such that  $\forall i \in N, \forall t_i \in T_i$ ,

$$U_i(g(\tilde{m})|t_i) \geq U_i(g(\tilde{m}_{-i}, m_i)|t_i) \quad \forall m_i : T_i \mapsto M_i.$$

Denote by  $B(G)$  the set of (mixed) Bayesian equilibria of the mechanism  $G$ . Let  $g(B(G))$  be the corresponding set of equilibrium outcomes.

The two definitions of implementability we shall consider in this paper are the following.

An SCS  $F$  is *exactly Bayesian implementable in mixed strategies* if there exists a mechanism  $G$  such that  $g(B(G)) \approx F$ .

An SCS  $F$  is *virtually Bayesian implementable in mixed strategies* if  $\forall \epsilon > 0$  there exists a mechanism  $G$  such that  $g(B(G)) \approx_\epsilon F$ .

A *direct mechanism* is one with  $M_i = T_i$  for all  $i \in N$ .

A *mixed deception* is a profile of functions,  $\alpha = (\alpha_i)_{i \in N}$ , where  $\alpha_i : T_i \mapsto \Delta(T_i)$  is  $\mathcal{T}_i$ -measurable,  $\alpha_i(t_i) \neq t_i$  for all  $t_i \in S_i$  for some set of types  $S_i$  of positive measure, for some  $i \in N$ . (Note that the identity function on  $T$  is not a deception.)

For an SCF  $f$  and a mixed deception  $\alpha$ ,  $f \circ \alpha$  denotes the SCF such that for each  $t \in T$ ,  $[f \circ \alpha](t) = f(\alpha(t))$ , where  $\alpha(t) = \prod_{i \in N} \alpha_i(t_i)$  denotes the measure on  $T$  induced by  $\alpha$  in state  $t$ , and  $f(\alpha(t))$  imposes  $f(t')$  on each of the realizations  $t'$  of the measure  $\alpha$ .

For an SCF  $f$ , a mixed deception  $\alpha$  and a type  $t_i \in T_i$ , let  $f_{\alpha_i(t_i)}(t') = f(t'_{-i}, \alpha_i(t_i))$  for all  $t' \in T$ , where  $f(t'_{-i}, \alpha_i(t_i))$  is defined as above.

The most fundamental condition in the theory of implementation with incomplete information using Bayesian equilibrium is (Bayesian) incentive compatibility:

An SCF  $f$  satisfies *incentive compatibility* if for all  $i \in N$ ,  $t_i \in T_i$  and all deceptions  $\alpha$ ,

$$U_i(f \mid t_i) \geq U_i(f_{\alpha_i(t_i)} \mid t_i).$$

An SCS  $F$  satisfies incentive compatibility if every SCF  $f \in F$  satisfies it.

In addition to incentive compatibility, the next two conditions are also necessary for exact Bayesian implementation in mixed strategies.

An SCS  $F$  satisfies *mixed Bayesian monotonicity* if for every  $f \in F$ , whenever it so happens that for any mixed deception  $\alpha$   $f \circ \alpha \notin F$ , there exist  $i \in N$ ,  $t_i \in T_i$  and an SCF  $y$  such that

$$U_i(y \circ \alpha \mid t_i) > U_i(f \circ \alpha \mid t_i) \text{ while } U_i(f \mid t'_i) \geq U_i(y_{\alpha_i(t_i)} \mid t'_i), \forall t'_i \in T_i. \quad (*)$$

If one imposes this condition to take account of only pure deceptions, one obtains the condition of Bayesian monotonicity (Jackson (1991)), already known in the literature. In Subsection 3.1 we shall provide a discussion that illustrates the differences between the two conditions.

Let  $E, E' \subset T$  be two common knowledge events among the agents in  $N$ . Denote by  $F(E)$  and  $F(E')$  the ranges assigned by the SCS  $F$  to  $E$  and  $E'$ , respectively, and denote by  $F(E, E')$  the range assigned by  $F$  to the concatenation of the two events.

An SCS  $F$  satisfies *closure* if  $F(E, E') = F(E) \times F(E')$ .

Jackson (1991) provides a characterization result for economic environments with at least three agents: he shows that an SCS is Bayesian implementable in pure strategies if and only if it is equivalent to an SCS that satisfies incentive compatibility, Bayesian monotonicity and closure. Our first task in this paper is to obtain a characterization of exact Bayesian implementation in mixed strategies.

### 3 Exact Implementation

In this section we show that closure, mixed Bayesian monotonicity and incentive compatibility are the only properties relevant to characterize exact Bayesian implementation in mixed strategies over economic environments.

We begin by introducing some additional pieces of notation.



For every SCF  $f \in F$  and for every mixed deception  $\alpha$  such that  $f \circ \alpha \notin F$ , a test-agent is any  $i \in N$  for whom condition (\*) holds. Denote by  $D_i(f)$  the set of mixed deceptions for which  $i$  is a test-agent at  $f$ . For each test-agent  $i$  and each mixed deception  $\alpha \in D_i(f)$ , fix an SCF  $y_i^{\alpha, f}$  satisfying (\*) for  $i$  of type  $t_i$ . Notice that condition (\*) concerns the SCF  $y$  only in those states in which agent  $i$  is of type  $\alpha_i(t_i)$ . There is, therefore, no loss of generality in assuming that  $y_i^{\alpha, f}$  is of the form:

$$y_i^{\alpha, f}(t_{-i}, t'_i) = y_i^{\alpha, f}(t_{-i}, t_i) \text{ for all } t_{-i} \in T_{-i} \text{ and } t'_i \in T_i.$$

Thus  $y_i^{\alpha, f}$  is constant over  $T_i$ .<sup>6</sup>

For each  $f \in F$  and each  $i \in N$ , let

$$C_i^f = \{f\} \cup \bigcup_{\alpha \in D_i(f)} \{y_i^{\alpha, f}\}.$$

Of course, it follows that if agent  $i$  is not a test-agent for any mixed deception  $\alpha$  at  $f \in F$ ,  $C_i^f = \{f\}$ .

For the remainder of the paper, we shall make two regularity assumptions on environments. First, we adapt the “no-total-indifference” (NTI) assumption made in Serrano and Vohra (2005) to our environments:

An environment  $\mathcal{E}$  satisfies *no-total-indifference* (NTI) if for every  $j \in N$ ,  $t_j \in T_j$  and  $T'_j \subseteq T_j$  such that  $\mu_j(T'_j | t_j) > 0$ , there exist  $a, a' \in A$  such that

$$\int_{T'_j} u_j(a, t) \mu_j(S_{-j} | t_j) dS_{-j} \neq \int_{T'_j} u_j(a', t) \mu_j(S_{-j} | t_j) dS_{-j}.$$

In addition, we shall make the *economic environment* assumption made in Jackson (1991). Prior to its definition, we shall define the concept of a splicing.

For two SCFs  $f$  and  $h$ , let the splicing of  $f$  and  $h$  be the following SCF:  $f/h_S$  defined as follows:

$$f/h_S(t) = \begin{cases} f(t) & \text{if } t \in S; \\ h(t) & \text{otherwise} \end{cases}$$

An environment  $\mathcal{E}$  is *economic* (E) if for any SCF  $f$  and any  $S \subseteq T^*$  with positive measure, there exist at least two agents  $i$  and  $j$ ,  $i \neq j$ , and SCFs  $h$

---

<sup>6</sup>This observation also applies to the definition of mixed Bayesian monotonicity, and to its virtual counterpart, found in the sequel.

and  $h'$  such that  $U_i(h'/f_S \mid t_i) > U_i(f \mid t_i)$  and  $U_j(h'/f_S \mid t_j) > U_j(f \mid t_j)$  for some  $t_i \in T_i$  and  $t_j \in T_j$ .

Our first result is the following characterization theorem for SCSs.

**Theorem 1** *Let  $n \geq 3$ . Suppose an environment  $\mathcal{E}$  satisfies NTI and E. Then, a social choice set  $F$  is exactly Bayesian implementable in mixed strategies if and only if it is equivalent to a social choice set that satisfies incentive compatibility, mixed Bayesian monotonicity and closure.*

### Proof of Theorem 1.

NECESSITY. Since the necessity of incentive compatibility and closure of an equivalent SCS is well known, we shall show that mixed Bayesian monotonicity is necessary for exactly implementing an incentive compatible SCS in mixed strategies.

Let  $f \in F$ , and consider a mixed deception  $\alpha$  such that  $f \circ \alpha \notin F$ . Suppose  $F$  is exactly implementable in mixed strategies. This implies that there exists a mechanism  $G$  and a strategy profile  $\sigma$  in it such that  $g(\sigma) \approx f$ . For  $\sigma \circ \alpha$  not to be a Bayesian equilibrium of the mechanism, a preference reversal as specified in the mixed Bayesian monotonicity condition must exist. Thus,  $F$  must satisfy mixed Bayesian monotonicity. ■

SUFFICIENCY. Suppose  $\hat{F} \approx F$  satisfies incentive compatibility, mixed Bayesian monotonicity and closure. We shall construct a canonical mechanism,  $G = ((M_i, \mathcal{M}_i)_{i \in N}, g)$  to exactly implement  $\hat{F}$  in mixed strategies.

Consider the following mechanism: The message set of agent  $i$  is defined as  $M_i = T_i \times F \times \prod_{f \in F} C_i^f \times \mathbf{F} \times I$ , where  $I$  is the set of non-negative integers, and recall that  $\mathbf{F}$  is the set of all SCFs. Let  $\mathcal{M}_i = \mathcal{T}_i \times \mathcal{F} \times \prod_f \mathcal{F} \times \mathcal{F} \times 2^I$  be its associated  $\sigma$ -algebra. Denote by  $(m_i^1, m_i^2, m_i^3, m_i^4, m_i^5)$  a typical message sent by agent  $i$ , by  $m_i^{3f}$  the part of the third component of agent  $i$  that is an element of  $C_i^f$ , and by  $m$  a profile of messages. Fixing a constant  $k \in (0, 1)$  and denoting by  $\bar{a}$  the uniform distribution over alternatives, the outcome function is defined by the following rules:

- (i) If there exists  $f \in \hat{F}$  such that at least  $n - 1$  agents announce  $m_i^2 = f$  and  $m_i^{3f} = f$ , then

$$g(m) = (1 - k)f(m^1) + \frac{k}{n} \sum_{i \in N} m_i^{3f}(m^1).$$

- (ii) Otherwise, denoting by  $j$  the agent with the lowest index among those who announce the highest integer,

$$g(m) = (1-k)m_j^2(m^1) + k\left[\frac{m_j^5}{m_j^5+1}m_j^4(m^1) + \frac{1}{n(m_j^5+1)}(\bar{a} + \sum_{i \neq j} m_i^2(m^1))\right].$$

To prove the theorem, we take the following steps:

STEP 1: Let  $f \in \hat{F}$ . A strategy profile where for each  $i \in N$  and each  $t_i \in T_i$ ,  $m_i(t_i) = (t_i, f, f, h, 0)$  is a Bayesian equilibrium of  $G$ . To see this, note that this strategy profile corresponds to the outcome of rule (i). Note that under this strategy profile, the outcome in state  $t$  is  $f(t)$ . Moreover, no unilateral deviation from it can trigger rule (ii), and therefore  $m_i^4$  or  $m_i^5$  have no effect on the outcome. Changing  $m_i^2$  has no effect on the outcome either. The only way an agent can change the outcome is by changing his announcement of  $m_i^1$  or  $m_i^3$ . Since  $f$  is incentive compatible, reporting a false type is not a profitable deviation for any agent. By condition (\*) it is not profitable to report a change in  $m_i^3$ . Nor is it possible to profit by changing both  $m_i^1$  and  $m_i^3$  because each  $y_i^{\alpha, f}$  in  $C_i^f$  is constant with respect to  $i$ 's type. Thus, as claimed, this profile is a Bayesian equilibrium of  $G$  and the outcome is  $f$ .

STEP 2. There cannot be an equilibrium  $\sigma$  that induces case (ii) with positive  $\sigma$ -measure over a positive-measure set of states. We argue by contradiction. Let  $T^I \subseteq T^*$  be such a positive measure set of states, i.e., those in which the integer game is induced with positive  $\sigma$ -measure (case (ii)). Let  $\tilde{n}$  be the supremum of the integers announced in any state in  $T^I$ , and let  $j$  be the lowest indexed agent who, without loss of generality, announces  $\tilde{n}$  in some state in  $T^I$ . Thus there exists a state  $t \in T^*$  in which the integer game is played, and is won by agent  $j$  of type  $t_j$  who announces the integer  $\tilde{n}$ . Let

$$T'_{-j} = \{t'_{-j} \in T_{-j} | (t_j, t'_{-j}) \in T^I\}.$$

By hypothesis, this set has positive measure. By construction, agent  $j$  of type  $t_j$ , by announcing  $n_j = \tilde{n}$ , wins the integer game in all states in  $\{t_j\} \times T'_{-j}$ . Since  $j$  is the lowest indexed agent who announces  $\tilde{n}$  in  $T^I$ , if agent  $j$  of type  $t_j$  changes her announcement of the integer to  $n'_j > \tilde{n}$ , everything else being the same, she continues to be the winner in (almost) the same states as before, namely  $\{t_j\} \times T'_{-j}$ . Let  $\sigma_j(t_j)$  include  $(\tilde{t}_j, \tilde{f}_j, (\tilde{z}_j^f)_{f \in F}, \tilde{h}_j, \tilde{n})$  on its support. Consider a measurable strategy  $\sigma'_j$  such that  $\sigma'_j(t_j) = (\tilde{t}_j, \tilde{f}_j, (\tilde{z}_j^f)_{f \in F}, h'_j, n'_j)$ ,

where  $n'_j > \tilde{n}$  and  $h'_j$  is chosen appropriately using NTI to yield a preferred lottery over these states where agent  $j$  of type  $t_j$  wins the integer game. This would increase type  $t_j$ 's expected utility, but this would contradict the hypothesis that the strategy profile  $\sigma$  is a Bayesian equilibrium.

STEP 3: There cannot be an equilibrium  $\sigma$  that, in a positive measure of states induces rule (i) of the outcome function  $g$  where, with positive  $\sigma$ -measure, exactly  $n - 1$  agents  $i$  coordinate on the announcement of the same  $f \in \hat{F}$  in their  $m_i^{3f}$ , while agent  $j$  announces  $m_j^{3f} = y_j^{\alpha, f} \neq f$ . Suppose this happened with positive  $\sigma$ -measure in a positive measure of states. Let type  $t_j$  of agent  $j$  be this “odd man out.” Then, in every state compatible with type  $t_j$  the outcome is

$$(1 - k)f(m^1) + k\left[\frac{n - 1}{n}f(m^1) + \frac{1}{n}y_j^{\alpha, f}(m^1)\right].$$

Then, by Assumption E, there exists a type  $t_i$  of agent  $i$  for whom this equilibrium outcome is not top ranked in his preference ordering. Then, any such type  $t_i$  can announce  $m_i^{3f} \neq f$ ,  $m_i^4$  and an integer  $m_i^5$  sufficiently high so that he becomes the winner in all states  $\{t_i\} \times T'_{-i}$  where types  $t_i$  and  $t_j$  are present. Note that for the marginal calculated by type  $t_i$  the set of these states has positive measure. Let  $\sigma_i(t_i)$  include  $(\tilde{t}_i, \tilde{f}_i, (\tilde{f}_i^f)_{f \in F}, \tilde{h}_i, \tilde{n}_i)$  in its support, and consider a measurable strategy  $\sigma'_i$  such that  $\sigma'_i(t_i) = (\tilde{t}_i, \tilde{f}_i, (z_i^f)_{f \in F}, h'_i, n'_i)$  where  $n'_i$  is chosen to ensure that  $i$  wins the integer game in all states in  $\{t_i\} \times T'_{-i}$ , and  $h'_i$  is chosen suitably to approximate a top ranked outcome. This implies a gain in terms of interim utility. But this contradicts the hypothesis that  $\sigma$  is a Bayesian equilibrium.

STEP 4: Finally, we claim that in any equilibrium of  $G$  under rule (i) where each agent  $i$  announces  $m_i^{3f} = f$  for some  $f \in \hat{F}$ , agents do not use a mixed deception  $\alpha$  where  $f \circ \alpha \notin \hat{F}$ . Suppose not, i.e., there is an equilibrium under rule (i) in which a mixed deception  $\alpha$  is used where  $f \circ \alpha \notin \hat{F}$ . Since  $\hat{F}$  satisfies mixed Bayesian monotonicity, there exists an agent  $i$  and an SCF  $y_i^{\alpha, f}$  satisfying (\*). Therefore, type  $t_i$  of agent  $i$  has an incentive to deviate and change only the third component of his announcement to  $m_i^{3f} = y_i^{\alpha, f}$ , which is a contradiction (of course, the messages of the other types of agent  $i$  can be changed to preserve the measurability of the strategy). Thus, either no mixed deception is used in equilibrium or the mixed deception being used is such that  $f \circ \alpha \in \hat{F}$ . In either case, therefore, the equilibrium outcome is in  $\hat{F}$ . This proves that our mechanism exactly implements  $\hat{F}$ . ■

As the reader may have noticed, Theorem 1 does not cover SCSs in the case of two agents. Note that if one tried to use the mechanism in the proof of Theorem 1 to implement an SCS with only two agents, one would run into the following problem. Suppose  $f$  and  $f'$  are two SCFs in the SCS  $F$  that we wish to implement. The mechanism allows for an equilibrium where both agents agree on the announcement of  $f$  as the second component of their strategy, and another in which they agree on  $f'$ . However, rule (i) ceases to be well defined if one of them announces  $f$ , while the other announces  $f'$ . It turns out that this is not just a failure of the particular mechanism employed to prove Theorem 1. Indeed, an extra condition is needed. Dutta and Sen (1994) study Bayesian implementation in pure strategies for the case of two agents, and prove that an extra condition having to do with the non-empty intersection of certain lower contour sets being non-empty is also necessary and sufficient. We would expect to find a similar condition to cover the case of mixed strategies.

However, the two-agent case is covered if one concentrates on SCFs, as shown in Theorem 2:

**Theorem 2** *Suppose an environment  $\mathcal{E}$  satisfies NTI and E. Then, a social choice function  $f$  is exactly Bayesian implementable in mixed strategies if and only if it is equivalent to a social choice function that satisfies incentive compatibility and mixed Bayesian monotonicity.*

### Proof of Theorem 2.

The necessity arguments are similar to those in Theorem 1, so we omit them (in particular, note how they are completely independent of the  $n \geq 3$  assumption, used only in the sufficiency part of Theorem 1). For sufficiency, the proof is also very similar, but we outline it: we construct a similar canonical mechanism to that in the proof of Theorem 1 with a simpler message set and  $\sigma$ -algebra for each agent  $i$ . Let  $\hat{f} \approx f$ . Then, agent  $i$ 's message set is  $M_i = T_i \times C_i^{\hat{f}} \times \mathbf{F} \times I$ . Let  $\mathcal{M}_i = \mathcal{T}_i \times \mathcal{F} \times \mathcal{F} \times 2^I$  be its associated  $\sigma$ -algebra. We modify the outcome function as follows:

- (i) If at least  $(n - 1)$  agents  $i$  announce  $m_i^2 = \hat{f}$ , then

$$g(m) = (1 - k)\hat{f}(m^1) + \frac{k}{n} \sum_{i \in N} m_i^2(m^1).$$

- (ii) Otherwise, denoting by  $j$  the agent with the lowest index among those who announce the highest integer,

$$g(m) = (1-k)m_j^2(m^1) + k\left[\frac{m_j^4}{m_j^4+1}m_j^3(m^1) + \frac{1}{n(m_j^4+1)}(\bar{a} + \sum_{i \neq j} m_i^2(m^1))\right].$$

With this mechanism, one can then follow the same steps as in the proof of Theorem 1 to show that it exactly implements  $\hat{f}$ . ■

### 3.1 Mixed Bayesian Monotonicity and Bayesian Monotonicity

In this subsection we shall discuss the differences between Bayesian monotonicity and its mixed counterpart. We first show that BM and MBM are different, by constructing an example that demonstrates that the mixed condition is more restrictive for SCSs.

**Example 1** Let  $N = \{1, 2\}$ , and  $A = \{a, b, c\}$ . Agent 1 is informed and his set of types is  $T = \{t, t'\}$ , and agent 2 is uninformed and comes in only one type. Preferences are as follows:

$$\begin{aligned} u_1((a, b, c), t) &= (0, 0, -1) & u_2((a, b, c), t) &= (2, 1, 0) \\ u_1((a, b, c), t') &= (0, 0, -1) & u_2((a, b, c), t') &= (1, 2, 1.5) \end{aligned}$$

Consider the following SCS

$$F \equiv \{(a, a), (a, b), (b, a), (b, b)\}.$$

Denote the four SCFs in  $F$  by  $f_1, \dots, f_4$ . Clearly, the SCS  $F$  is fully implemented with a mechanism in which only agent 1 sends messages, his message set is  $F \times T$  with typical message  $(\hat{f}, \hat{t})$ , and the outcome is  $\hat{f}(\hat{t})$ . Thus,  $F$  satisfies Bayesian monotonicity.

However,  $F$  does not satisfy MBM. Consider the SCF  $f_2 = (a, b) \in F$  and the mixed deception  $\alpha$  in which each type of agent 1 randomizes equally between telling the truth and lying. It is clear that the resulting SCF,  $f_2 \circ \alpha$ , which imposes in each state the random alternative  $(1/2)a + (1/2)b$ , is not in  $F$ . Therefore, one needs to find a test-agent.

Agent 2 cannot be a test agent: we know that  $U_2(f_2 \circ \alpha) = 3/2$ . Any SCF  $y$  is such that for this deception  $\alpha$  the random alternative  $(1/2)y(t) + (1/2)y(t')$

is implemented. Denote this probability distribution by  $(\lambda, \mu, 1 - \lambda - \mu)$ . Then, his expected utility from  $y \circ \alpha$  is:

$$U_2(y \circ \alpha) = (3/2)\lambda + (3/2)\mu + (3/4)(1 - \lambda - \mu) \leq 3/2,$$

so agent 2 cannot be used as a test-agent.

Clearly, no type of agent 1 can be used as a test-agent either. Using the same argument as above for agent 2, note that  $U_1(f_2, s) = U_1(f_2 \circ \alpha, s) = 0$  for  $s = t, t'$ . And

$$U_1(y \circ \alpha, s) = \lambda + \mu - 1 \leq 0.$$

An SCS  $F$  satisfies the *convex range property* if, whenever it is true that for  $f \in F$  and for a collection of pure deceptions  $\beta$ , one has that  $f \circ \beta \in F$ , then it is true that for every mixed deception  $\alpha$  whose support are the  $\beta$ 's, it is the case that  $f \circ \alpha \in F$ .

Note that every SCF trivially satisfies this property. More generally, we have the following result:

**Proposition 1** *Mixed Bayesian monotonicity implies Bayesian monotonicity. Conversely, Bayesian monotonicity and the convex range property imply mixed Bayesian monotonicity.*

**Proof of Proposition 1.** By definition, it is clear that MBM implies BM because it imposes the requirement of the appropriate preference reversal over all mixed deceptions, which include pure deceptions as particular cases.

We now show that BM and the convex range property imply MBM. Consider a non-pure deception  $\alpha$  such that  $f \circ \alpha \notin F$ . Let  $\beta$  be a pure deception in the support of  $\alpha$  satisfying that  $f \circ \beta \notin F$ . The existence of such a  $\beta$  is guaranteed by the convex range property.

Since  $\beta$  is a pure deception and  $F$  satisfies BM, there exist  $i \in N$ ,  $t_i \in T_i$  and an SCF  $y$  such that

$$U_i(y \circ \beta \mid t_i) > U_i(f \circ \beta \mid t_i) \text{ while } U_i(f \mid t'_i) \geq U_i(y_{\beta_i(t_i)} \mid t'_i), \forall t'_i \in T_i.$$

Since  $\alpha$  can be expressed as a function of  $\beta$  and the other pure deceptions in its support, we choose  $f$  on the lhs and rhs of (\*) as the pair of functions associated with each of the other pure deceptions. Then, we get a preference reversal as in (\*) using  $\alpha$  on the basis only of the inequalities just written for  $\beta$ . ■

## 4 Approximate Implementation

This section will follow parallel steps to the previous one. We will be able to prove our results for virtual implementation by appealing directly to our results for exact implementation.

We begin with a new definition:

An SCF  $F$  satisfies *mixed virtual monotonicity* if for every  $f \in F$ , there exists an incentive compatible SCF  $x$  such that whenever happens that for any mixed deception  $\alpha$ ,  $f \circ \alpha \notin F$ , there exists  $i \in N$ ,  $t_i \in T_i$  and an SCF  $y$  such that

$$U_i(y \circ \alpha \mid t_i) > U_i(x \circ \alpha \mid t_i) \text{ while } U_i(x \mid t'_i) \geq U_i(y_{\alpha_i(t_i)} \mid t'_i), \forall t'_i \in T_i. \quad (**)$$

The main difference between this condition and mixed Bayesian monotonicity is that the preference reversal in the new condition does not necessarily involve an SCF  $f \in F$ , but some other incentive compatible SCF  $x$  in the environment. Note how mixed virtual monotonicity differs from the virtual monotonicity condition proposed in Serrano and Vohra (2005) in several respects. To begin with, in that paper the condition was formulated only for SCFs. In addition, mixed virtual monotonicity is a strengthening of virtual monotonicity for two reasons: first, it is imposed on mixed deceptions that undermine  $F$ , instead of simply on pure deceptions; and second, there is a change in the quantifiers of the condition: in virtual monotonicity, for each deception, there is an incentive compatible  $x$  and these may differ across deceptions (see Serrano and Vohra (2005, Section 3)). On the other hand, it can be shown that the stronger condition with the order of quantifiers reversed can replace the one used in Serrano and Vohra (2005) to show their result. The reason is that, although it is a stronger condition, it is also necessary for virtual Bayesian implementation in pure strategies.

The version of virtual monotonicity in Serrano and Vohra (2005) is used in the sufficiency part of their theorem to construct a canonical mechanism that implements an SCF that puts probability  $(1 - \epsilon)$  on  $f$  and probability  $\epsilon$  on a convex combination of all the SCFs  $x$  that are associated with each pure deception. Using the extension of that same condition to mixed deceptions would have been insufficient to construct a similar canonical mechanism for our purpose here, since we would end up with an uncountable sum of SCFs. The version proposed here solves this problem, since a single  $x$  is associated with each mixed deception that undermines each SCF of interest. In any event, we shall provide a shortcut to the sufficiency proof and not rely on the



construction of a new canonical mechanism. This is the advantage of having developed the exact implementation results in the previous section (in the end, virtual or approximate implementation is simply the exact implementation of a near-by SCS).

Our next result is the following characterization theorem for SCSs in environments with at least three agents.

**Theorem 3** *Let  $n \geq 3$ . Suppose an environment  $\mathcal{E}$  satisfies NTI and E. Then, a social choice set  $F$  is virtually Bayesian implementable in mixed strategies if and only if it is equivalent to a social choice set that satisfies incentive compatibility, mixed virtual monotonicity and closure.*

### Proof of Theorem 3.

NECESSITY. Since the necessity of incentive compatibility and closure are standard, we shall show that mixed virtual monotonicity is necessary for virtually implementing in mixed strategies an incentive compatible SCS.

Suppose then that  $F$  is virtually Bayesian implementable in mixed strategies, i.e., for every  $\epsilon > 0$  there exists an SCS  $F^\epsilon$  that is exactly implementable in mixed strategies and is  $\epsilon$ -approximate to  $F$ . Let  $f \in F$  and call  $f^\epsilon = \pi(f)$ , where  $\pi$  is the bijection between  $F$  and  $F^\epsilon$ . Consider an arbitrary mixed deception  $\alpha$  such that  $f \circ \alpha \notin F$ . Choosing  $\epsilon > 0$  small enough, one has that  $f^\epsilon \circ \alpha \notin F^\epsilon$ . It follows from Theorem 1 that  $F^\epsilon$  satisfies incentive compatibility and mixed Bayesian monotonicity. Therefore,  $f^\epsilon$  satisfies the required preference reversal in (\*\*) for any such mixed deception  $\alpha$ . That is,  $f^\epsilon$  is incentive compatible and is such that for every  $\alpha$  satisfying that  $f \circ \alpha \notin F$ , there exists  $i \in N$ ,  $t_i \in T_i$  and an SCF  $y$  such that

$$U_i(y \circ \alpha \mid t_i) > U_i(f^\epsilon \circ \alpha \mid t_i) \text{ while } U_i(f^\epsilon \mid t'_i) \geq U_i(y_{\alpha_i(t_i)} \mid t'_i), \forall t'_i \in T_i. \quad (1)$$

But this means that  $F$  satisfies mixed virtual monotonicity. ■

SUFFICIENCY. Suppose  $\hat{F} \approx F$  satisfies incentive compatibility, mixed virtual monotonicity and closure. The novelty of this proof is that, rather than constructing a canonical mechanism to do the job, we shall rely on the properties assumed and Theorem 1 to obtain a direct proof. We shall show that for every  $\epsilon > 0$ , there exists  $F^\epsilon$  such that  $\hat{F} \approx_\epsilon F^\epsilon$  and satisfying mixed Bayesian monotonicity.

Suppose  $\hat{F}$  satisfies mixed virtual monotonicity, incentive compatibility and closure. Let  $f \in \hat{F}$ . By mixed virtual monotonicity, there exists an incentive compatible SCF  $x$  and an SCF  $y$  exhibiting the appropriate preference reversal as in (\*\*) for every  $\alpha$  satisfying  $f \circ \alpha \notin \hat{F}$ . Let  $f^\epsilon = (1 - \epsilon)f + \epsilon x$  and  $y' = (1 - \epsilon)f + \epsilon y$ . Note that, by incentive compatibility of  $\hat{F}$ , the SCF  $f^\epsilon$  is also incentive compatible.

Define the SCS  $F^\epsilon$  to be the following set:

$$F^\epsilon \equiv \{f^\epsilon : f \in \hat{F}\}.$$

For  $\epsilon > 0$  small enough, we claim that  $F^\epsilon$  satisfies mixed Bayesian monotonicity. Since

$$U_i(y' \circ \alpha | t_i) - U_i(f^\epsilon \circ \alpha | t_i) = \epsilon[U_i(y \circ \alpha | t_i) - U_i(x \circ \alpha | t_i)],$$

it follows from (\*\*) that

$$U_i(y' \circ \alpha | t_i) > U_i(f^\epsilon \circ \alpha | t_i) \tag{2}$$

From (\*\*) we also know that

$$U_i(x | t'_i) \geq U_i(y_{\alpha_i(t_i)} | t'_i), \forall t'_i \in T_i.$$

Thus

$$(1 - \epsilon)U_i(f | t'_i) + \epsilon U_i(x | t'_i) \geq (1 - \epsilon)U_i(f | t'_i) + \epsilon U_i(y_{\alpha_i(t_i)} | t'_i), \forall t'_i \in T_i.$$

Since,  $f$  is incentive compatible,

$$U_i(f | t'_i) \geq U_i(f_{\alpha_i(t_i)} | t'_i), \forall t'_i \in T_i$$

The last two inequalities imply that

$$U_i(f^\epsilon | t'_i) \geq U_i(y'_{\alpha_i(t_i)} | t'_i), \forall t'_i \in T_i. \tag{3}$$

So the appropriate reversal happens for every mixed deception  $\alpha$  such that  $f \circ \alpha \notin \hat{F}$ . However, for  $\epsilon > 0$  small enough, this set of mixed deceptions is equivalent to the set satisfying that  $f^\epsilon \circ \alpha \notin F^\epsilon$ . Therefore, since all this is true for every  $f^\epsilon \in F^\epsilon$ , it follows that  $F^\epsilon$  satisfies mixed Bayesian monotonicity.

It is also true that the SCS  $F^\epsilon$  defined above satisfies closure because so does  $\hat{F}$ .

Therefore, since for every  $\epsilon > 0$  sufficiently small,  $F^\epsilon$  satisfies incentive compatibility, mixed Bayesian monotonicity and closure, and by construction  $F^\epsilon \approx_\epsilon \hat{F}$ , it follows from Theorem 1 that  $\hat{F}$  is virtually implementable in mixed strategies. ■

Following parallel steps of reasoning, we obtain the following characterization of SCFs that are virtually implementable in mixed strategies.

**Theorem 4** *Suppose an environment  $\mathcal{E}$  satisfies NTI and E. Then, a social choice function  $f$  is virtually Bayesian implementable in mixed strategies if and only if it is equivalent to a social choice function that satisfies incentive compatibility and mixed virtual monotonicity.*

The proof of Theorem 4 is similar to that of Theorem 3, but it uses Theorem 2 instead of Theorem 1. Also, the relationship between virtual monotonicity and mixed virtual monotonicity is similar to that between Bayesian monotonicity and its mixed counterpart (see Subsection 3.1). In particular, for SCFs, mixed virtual monotonicity is not more restrictive than the condition that uses only pure deceptions.

## 5 Discussion

This section discusses a couple of related approaches in the literature. First, we relate our work to Maskin’s (1999) treatment of Nash implementation in mixed strategies. And second, we shall draw connections between our conditions and the condition of measurability, introduced in Abreu and Matsushima (1992b) – we shall refer to this condition as A-M measurability. The latter will have implications on the implementing mechanisms used in the current paper.

### 5.1 Maskin’s Approach

Maskin (1999) is the classic paper on Nash implementation. The appendix to the paper extends the analysis to cover mixed strategies (see also Maskin and Sjöström (2002, Section 4.3)). Maskin’s approach differs from ours in that, consistent with most of the Nash implementation literature, the social

choice rule assigns pure alternatives to each ex-post preference profile, instead of probability distributions over alternatives. The notion of implementability that he proposes requires that each realization of a mixed-strategy Nash equilibrium be consistent with the social choice rule. This “ex-post” approach allows him to dispense with any specific assumptions concerning risk preferences. In this framework, he shows that his standard monotonicity condition, which is necessary for Nash implementability, is also sufficient along with no-veto power if there are at least three agents.<sup>7</sup>

Especially when one includes incomplete information in the model, it makes sense to allow that social choice functions map into the probability simplex over alternatives (e.g., optimal risk sharing in a contract will lead to different lotteries, depending on the parties’ risk preferences), and this is what we do. When one restricts attention to complete information environments and non-random social choice rules, our approach reduces to Maskin’s. That is, even though our implementability requirement is different (i.e., that every equilibrium outcome –pure or mixed– agree with the rule of interest), if there is complete information among the agents and randomness plays no role in the rule, one should expect that the two approaches relate to one another. Indeed, we argue now that mixed Bayesian monotonicity reduces to Maskin monotonicity in this case.

Suppose that  $E$  and  $E'$  are two complete information environments. Let  $a \in A$  be an alternative assigned by the social choice correspondence when the environment is  $E$ , but not assigned when the environment is  $E'$ . Denoting the type profiles in  $E$  and  $E'$  by  $t$  and  $t'$ , respectively, Maskin monotonicity requires the existence of an agent  $i$  and an alternative  $y$  satisfying that  $u_i(a, t_i) \geq u_i(y, t_i)$ , and  $u_i(y, t'_i) > u_i(a, t'_i)$ .

Now in our notation, since the social choice correspondence will assign to environment  $E'$  some other alternative  $b \in A$  and recalling that our social choice sets satisfy closure, it means that  $a/b_E \in F(E, E')$ , while  $a/a_E \notin F(E, E')$ . The existence of the agent  $i$  and the alternative  $y$  of the previous paragraph imply that for a pure deception satisfying  $\alpha(t) = t'$  and for the SCF  $\tilde{y}$  such that  $\tilde{y}(t) = \tilde{y}(t') = y$ ,  $f(t'') = \tilde{y}(t'')$  for all other  $t''$ , we have  $u_i(f(t), t''_i) \geq u_i(\tilde{y}, t''_i)$  for all types  $t''_i$ . However,  $u_i(\tilde{y} \circ \alpha, t'_i) > u_i(f \circ \alpha, t'_i)$ . Thus, a preference reversal in the sense of Maskin can always be associated

---

<sup>7</sup>In a framework similar to ours, making use of random mechanisms, Benoit and Ok(2008) and Bochet (2007) investigate conditions under which no-veto power can be dispensed with in Maskin’s theorem. Since we confine our attention to economic environments, the issue is moot for us.

with a reversal in our sense, if one uses a pure deception like this.

Thus, the conditions of Maskin monotonicity and Bayesian monotonicity coincide for this case. They differ if one takes into account non-pure deceptions. That is, if a pair of alternatives  $a, b \in A$  is selected by the correspondence in two complete information states  $E$  and  $E'$ , mixed Bayesian monotonicity will require that the social choice set also include the probability distributions over  $a$  and  $b$  in those states. Of course, if the model rules out random outcomes, this difference will be unimportant for such a model. And if one considers social choice functions instead of multi-valued correspondences, there is no difference.

By the same token, continuing with our restriction to complete information environments, one can see that the condition of mixed virtual monotonicity is trivial when applied over functions, and only adds the restrictions coming from the convex range property when one considers correspondences. This relates to the very permissive results in Abreu and Sen (1991) and Matsushima (1988) for virtual Nash implementation.

## 5.2 Mixed Virtual Monotonicity and A-M Measurability

For this subsection we shall follow the assumptions made in most of the virtual implementation papers and confine our discussion to SCFs. Abreu and Matsushima (1992b) study virtual implementation with incomplete information employing the solution concept of iterative elimination of strictly dominated strategies. Relying on a numeraire to implement small punishments out of equilibrium, they show that virtual implementation in their sense is characterized by incentive compatibility and A-M measurability; the reader is referred to Abreu and Matsushima (1992b) or to Duggan (1997) for the definition of A-M measurability. Obviously in this case, since any implementation in iterative elimination of strictly dominated strategies implies implementation in mixed-strategy equilibrium, it follows that, for incentive compatible SCFs, measurability implies virtual monotonicity.<sup>8</sup> There is a sense in which all these conditions are rather weak. Under type diversity (see Serrano and Vohra (2005), and its references to previous sources), every SCF

---

<sup>8</sup>Duggan (1997) provides incentive consistency as a sufficient condition, together with incentive compatibility, for virtual implementation in mixed-strategy Bayesian equilibrium. His condition is not necessary, however, as shown in Serrano and Vohra (2001).

satisfies A-M measurability, incentive consistency and virtual monotonicity. If there are at least three alternatives, “almost every” environment satisfies type diversity.

To illustrate the gap between A-M measurability and implementation in Bayesian equilibrium, Duggan (1997) provides the following example. We elaborate further on it:

**Example 2** *Consider the following environment, in which  $N = \{1, 2\}$ . Each agent comes in two types:  $T_1 = \{t_1, t'_1\}$  and  $T_2 = \{t_2, t'_2\}$ . The set of states is the Cartesian product of each agent's set of types, and each state is equally likely. There are four alternatives:  $A = \{x_1, x_2, x_3, x_4\}$ . Agents' ex-post preferences in each state are as follows:*

$$u_1((x_1, x_2, x_3, x_4), (t_1, t_2)) = (3, 2, 2.5, 2), u_2((x_1, x_2, x_3, x_4), (t_1, t_2)) = (2, 3, 2, 2.5);$$

$$u_1((x_1, x_2, x_3, x_4), (t_1, t'_2)) = (3, 2, 3, 2), u_2((x_1, x_2, x_3, x_4), (t_1, t'_2)) = (2, 3, 2, 2.5);$$

$$u_1((x_1, x_2, x_3, x_4), (t'_1, t_2)) = (3, 2, 2.5, 2), u_2((x_1, x_2, x_3, x_4), (t'_1, t_2)) = (2, 3, 2, 3);$$

$$u_1((x_1, x_2, x_3, x_4), (t'_1, t'_2)) = (3, 2, 3, 2), u_2((x_1, x_2, x_3, x_4), (t'_1, t'_2)) = (2, 3, 2, 3).$$

*As argued in Duggan (1997), only constant SCFs satisfy A-M measurability in this environment. Duggan (1997) considers the following SCF:*

$$f(t_1, t_2) = x_1, \quad f(t_1, t'_2) = x_2, f(t'_1, t_2) = x_4, \quad f(t'_1, t'_2) = x_3,$$

*and shows it to be exactly Bayesian implementable in pure strategies via its direct mechanism.*

*Therefore, it follows that such an SCF satisfies Bayesian monotonicity. By our observation preceeding Proposition 1, it also satisfies mixed Bayesian monotonicity, and thus, by our Theorem 2, it is also exactly implementable in mixed-strategy Bayesian equilibrium.*

*However, the implementing mechanism that yields implementation in mixed strategies will not be the direct mechanism: indeed, as the reader can check,*

*the following deception, which undermines  $f$ , is a mixed strategy equilibrium of the direct mechanism for  $f$ :*

$$\begin{aligned}\alpha_1(t_1) &= (1/3, 2/3), & \alpha_1(t'_1) &= (1/2, 1/2); \\ \alpha_2(t_2) &= (1/3, 2/3) & \alpha_2(t'_2) &= (1/2, 1/2).\end{aligned}$$

Abreu and Matsushima (1992b) also show that A-M measurability is necessary for implementation in Bayesian equilibrium when the implementing mechanism is regular (this condition amounts to the property that best replies always exist, thereby ruling out devices like integer games). It follows that in the example just described, due to Duggan, the SCF  $f$ , which is not virtually implementable in iteratively undominated strategies because it is not A-M measurable, must be implemented by means of a mechanism in which non-regular features are essential (e.g., the mechanism of the proof of Theorem 1, properly modified as described in the proof of Theorem 2). Of course, this observation is general: those SCFs that can be decentralized using equilibrium behavior, and cannot using iteratively undominated strategies, necessitate non-regular mechanisms.

## References

- Abreu, D. and H. Matsushima (1992a). Virtual Implementation in Iteratively Undominated Strategies: Complete Information, *Econometrica*, 60, 993-1008.
- Abreu, D. and H. Matsushima (1992b). Virtual Implementation in Iteratively Undominated Strategies: Incomplete Information, Unpublished Manuscript, Princeton University.
- Abreu, D. and A. Sen (1991). Virtual Implementation in Nash Equilibrium, *Econometrica*, 59, 997-1021.
- Artemov, G., T. Kunimoto and R. Serrano (2007). Robust Virtual Implementation with Incomplete Information: Towards a Reinterpretation of the Wilson Doctrine, Economics Working Paper 2007-06, Brown University.
- Benoit, J.-P. and E. A. Ok (2008). Nash Implementation without No-Veto Power, *Games and Economic Behavior*, 64, 51-67.
- Bergemann, D. and S. Morris (2009). Robust Virtual Implementation, *Theoretical Economics*, 4, 45-88.

- Bochet, O. (2007). Nash Implementation with Lottery Mechanisms. *Social Choice and Welfare*, 28, 111-125.
- Duggan, J. (1997). Virtual Bayesian Implementation, *Econometrica*, 65, 1175-1199.
- Dutta, B. and A. Sen (1994). Two-Person Bayesian Implementation, *Economic Design*, 1, 41-54.
- Jackson, M. O. (1991). Bayesian Implementation, *Econometrica*, 59, 461-477.
- Jackson, M. O. (1992). Implementation in Undominated Strategies: A Look at Bounded Mechanisms, *Review of Economic Studies*, 59, 757-775.
- Maskin, E. S. (1999). Nash Equilibrium and Welfare Optimality, *Review of Economic Studies*, 66, 23-38.
- Maskin, E. S. and T. Sjöström (2002). Implementation Theory, in K. J. Arrow, A. Sen and K. Suzumura (eds.) *Handbook of Social Choice and Welfare* (vol. I), New York, Elsevier Science B.V.
- Matsushima, H. (1988). A New Approach to the Implementation Problem, *Journal of Economic Theory* 45, 128-144.
- Mookherjee, D. and S. Reichelstein (1990). Implementation via Augmented Revelation Mechanisms, *Review of Economic Studies*, 57, 453-475.
- Palfrey, T. R. and S. Srivastava (1987). On Bayesian Implementable Allocations, *Review of Economic Studies*, 54, 193-208.
- Palfrey, T. R. and S. Srivastava (1989). Implementation with Incomplete Information in Exchange Economies, *Econometrica*, 57, 115-134.
- Postlewaite, A. and D. Schmeidler (1986). Implementation in Differential Information Economies, *Journal of Economic Theory*, 39, 14-33.
- Serrano, R. and R. Vohra (2001). Some Limitations of Virtual Bayesian Implementation, *Econometrica*, 69, 785-792.
- Serrano, R. and R. Vohra (2005). A Characterization of Virtual Bayesian Implementation, *Games and Economic Behavior*, 50, 312-331.