

NASH IMPLEMENTATION USING SIMPLE MECHANISMS WITHOUT UNDESIRABLE MIXED-STRATEGY EQUILIBRIA

MARIA GOLTSMAN

ABSTRACT. This note shows that, in separable environments, any monotonic social choice function can be Nash implemented with a mechanism that makes no use of integer games or similar constructions and admits no mixed-strategy Nash equilibria in undominated strategies that lead to undesirable outcomes with positive probability.

Implementation theory studies the problem of designing a mechanism such that all its equilibria yield outcomes coinciding with the value of a given social choice correspondence. In a seminal paper, Maskin (1999) provides a necessary and almost sufficient condition (monotonicity) for a social choice correspondence to be implementable in Nash equilibrium. The sufficiency proof is by construction, and the constructed mechanism makes use of an integer game, whereby each player has to name an integer, and the one who names the highest integer gets to choose the allocation to be implemented. The integer game is helpful in eliminating undesired equilibria, but the best-response correspondences are not well-defined in this game, so it is not clear how a participant should act in response to certain strategies of his opponents. Saijo (1988) proposes an alternative proof of Maskin's result, which replaces the integer game with a modulo game. The modulo game involves each agent naming an integer from a bounded set; the agent whose index equals the sum of the named integers modulo the number of agents gets to choose the allocation. The modulo game does not require an infinite strategy space, but can possess mixed-strategy equilibria that can result in undesired outcomes with positive probability.

A number of papers (for example, Yamato (1993), Sjöström (1996), Duggan (2003)) have proved that in certain environments, it is possible to construct mechanisms for Nash implementation that make no use of integer or modulo games. However, these papers share another drawback: they look only at pure-strategy Nash equilibria and ignore the potential existence of undesirable mixed-strategy Nash equilibria. At the same time, Jackson (1992)

Date: June 2011.

This note is based on Chapter 3 of my dissertation submitted to Northwestern University. I am grateful to Kim-Sau Chung, Gregory Pavlov, Asher Wolinsky and Charles Zheng for helpful suggestions. All errors are mine.

shows that ignoring mixed-strategy equilibria can be dangerous by providing an example of a mechanism that Nash implements a given social choice function, but has a mixed-strategy Nash equilibrium that Pareto dominates the pure-strategy one.

It is known that implementation in equilibrium concepts other than Nash equilibrium allows for mechanisms that do not employ integer games or have undesirable mixed-strategy equilibria (see Abreu and Matsushima (1992a,b) and Arya et al. (1995) for virtual implementation, Jackson et al. (1994) and Sjöström (1994) for implementation in undominated Nash equilibrium, and Abreu and Matsushima (1994) for implementation via iterative deletion of weakly dominated strategies). Duggan and Roberts (2002) exhibit a mechanism that has these properties in the Nash implementation framework; however, this paper considers a very specific problem (the allocation of pollution rights) and looks only at the efficient social choice correspondence.

This note shows that in a certain class of economic environments, any monotonic social choice function can be Nash implemented with a mechanism that makes no use of integer or modulo games; moreover, any strategy profile that survives one round of elimination of weakly dominated strategies and one round of elimination of strictly dominated strategies results in the desired outcome.¹ Therefore, the mechanism does not have any mixed-strategy Nash equilibria in undominated strategies that result in undesirable outcomes with positive probability. The class of environments considered includes, for example, pure exchange economies and public good provision problems with quasilinear utilities where sufficiently high fines can be collected from the agents.

The papers most closely related to this note are Jackson et al. (1994) and Sjöström (1994). These papers show that in economic environments any social choice function can be implemented in undominated Nash equilibrium using a mechanism without integer or modulo games; moreover, their mechanisms do not admit undesirable mixed-strategy Nash equilibria in undominated strategies (however, undesirable pure-strategy Nash equilibria are possible). The mechanisms that these papers use are similar to the one used in the present note.

The note is also related to the literature on double implementation. The double implementation exercise involves choosing two different equilibrium concepts and designing a mechanism such that any equilibrium according to either concept results in a desirable outcome. (Thus the present paper provides a mechanism for double implementation in Nash

¹As a consequence, in every state, all strategy profiles in the set $S^\infty W$ (see Dekel and Fudenberg (1990) and Börgers (1994)) result in the desired outcome.

equilibrium and strategies that survive one round of elimination of weakly dominated strategies and one round of elimination of strictly dominated strategies). Yamato (1993) studies double implementation in Nash equilibrium and undominated Nash equilibrium in economic environments; Mizukami and Watayama (2007) and Saijo et al. (2007) look at double implementation in Nash equilibrium and dominant strategies.

Finally, a recent paper by Mezzetti and Renou (2010) studies the problem of mixed Nash implementation in general environments. A social choice correspondence is mixed Nash implementable if for any outcome in its image, there is a Nash equilibrium where this outcome arises with positive probability, and no Nash equilibrium can result in an undesirable outcome with positive probability.² Among other results, this paper provides sufficient conditions for a social choice correspondence to be mixed Nash implementable with a bounded mechanism in separable environments. These conditions, as well as the definition of a separable environment, are different from those in the present note; however, the biggest difference is that Mezzetti and Renou (2010) allow for mechanisms that use lotteries, while the present note does not.

1. THE SETUP

Consider an economy with a finite number $N \geq 2$ of agents (with a slight abuse of notation, we will also denote the set of agents by N). There is a nonempty set \mathcal{A} of social alternatives, and the preferences of agent i over the alternatives in \mathcal{A} are determined by his type θ_i . Let $\Theta = \times_{i \in N} \Theta_i$, where Θ_i is the set of possible types of agent i ; we will assume that for every $i \in N$, Θ_i is finite. Let $u_i : \Theta_i \times \mathcal{A} \rightarrow \mathcal{R}$ be the utility function of type θ_i of agent i . An *environment* is a triple $\mathcal{E} = (\mathcal{A}, \Theta, (u_i)_{i \in N})$.³

A *social choice function* (SCF) f maps Θ to \mathcal{A} and is interpreted as the social alternative that is “desirable” if the type profile is θ . A *mechanism* is a tuple $\Gamma = ((S_i)_{i \in N}, g)$, where S_i are arbitrary sets, and g is a function:

$$g : S_1 \times \cdots \times S_N \rightarrow \mathcal{A}$$

²Serrano and Vohra (2010) study mixed Bayesian-Nash implementation in environments with incomplete information.

³Note that we assume that the state space has a product structure, i.e. any type profile $\theta \in \times_{i \in N} \Theta_i$ can occur. This assumption is often made in economic environments: see, e.g., Yamato (1993), Sjöström (1994, 1996, 1999), Mizukami and Wakayama (2007), Saijo et al. (2007), and Section 5 of Moore and Repullo (1988); Saijo (1988) makes this assumption in a more general environment.

A mechanism Γ induces a game, where the agents are the players, the type profile θ is common knowledge, the action set for agent i is S_i , and the payoff of agent i of type θ_i when a strategy profile s is played equals $u_i(\theta_i, g(s))$.

A mechanism Γ *implements SCF f in Nash equilibrium* if:

$$(1) \quad \forall \theta \in \Theta, \exists s \in S_1 \times \dots \times S_N : g(s) = f(s)$$

$$\text{and } \forall i \in N, \forall s'_i \neq s_i, u_i(\theta_i, g(s)) \geq u_i(\theta_i, g(s'_i, s_{-i}))$$

and

$$(2) \quad \forall \theta \in \Theta, s \in S_1 \times \dots \times S_N,$$

$$\text{if } \forall i \in N, \forall s'_i \neq s_i, u_i(\theta_i, g(s)) \geq u_i(\theta_i, g(s'_i, s_{-i})), \text{ then } g(s) = f(\theta).$$

Condition (1) says that for any state θ , there is a Nash equilibrium in the game induced by Γ which results in the outcome $f(\theta)$. Condition (2) says that any pure strategy Nash equilibrium outcome of the game induced by Γ has to be $f(\theta)$.

A SCF f is *implementable in Nash equilibrium* if there exists a mechanism Γ that implements f in Nash equilibrium.

Definition 1. A SCF f is **monotonic** if for any $\theta, \theta' \in \Theta$ such that $f(\theta) \neq f(\theta')$, there exist $i \in N$ (a **test agent**) and $b \in \mathcal{A}$ (a **test allocation**) such that $u_i(\theta_i, f(\theta)) \geq u_i(\theta_i, b)$ and $u_i(\theta'_i, b) > u_i(\theta'_i, f(\theta))$.

Maskin (1999) proves that any Nash implementable SCF is monotonic, and if $N \geq 3$, then any monotonic SCF satisfying an additional condition (no veto power) is Nash implementable.

2. SEPARABLE ENVIRONMENTS

In this section, we define a separable environment and give some examples.

Definition 2. Environment \mathcal{E} is **separable** relative to a SCF f if it satisfies the following conditions:

$$(3) \quad \exists \underline{w}, \bar{w} : \forall a \in f(\Theta), \forall i \in N, \forall \theta_i \in \Theta_i, u_i(\theta_i, \underline{w}) < u_i(\theta_i, \bar{w}) < u_i(\theta_i, a)$$

$$(4) \quad \forall a \in \mathcal{A}, \forall N_1, N_2 \subseteq N : N_1 \cap N_2 = \emptyset, \exists [a]^{N_1, N_2} \in \mathcal{A} :$$

$$\forall \theta \in \Theta, \begin{cases} u_j(\theta_j, [a]^{N_1, N_2}) = u_j(\theta_j, a), \forall j \in N_1; \\ u_j(\theta_j, [a]^{N_1, N_2}) = u_j(\theta_j, \underline{w}), \forall j \in N_2; \\ u_j(\theta_j, [a]^{N_1, N_2}) = u_j(\theta_j, \bar{w}), \forall j \in N \setminus (N_1 \cup N_2). \end{cases}$$

$$(5) \quad \forall i \in N, \theta_i, \theta'_i \in \Theta_i, \theta_i \neq \theta'_i, \exists a, b \in \mathcal{A} :$$

$$\begin{cases} u_i(\theta_i, a) \geq u_i(\theta_i, b) \geq u_i(\theta_i, \bar{w}), \\ u_i(\theta'_i, b) > u_i(\theta'_i, a) \geq u_i(\theta'_i, \bar{w}) \end{cases}$$

Condition (3) says that there exist at least two outcomes, \underline{w} and \bar{w} , that are ranked in the same way for any type of any agent and are worse than any outcome in $f(\Theta)$. In quasilinear setting, these outcomes can be thought of two different levels of fines that are collected from the agents. The outcomes will serve as punishments in the mechanism. Condition (4) says that it is feasible to impose punishments on a subset of agents without affecting the agents outside the subset. Our definition is closely related to the definition of a separable environment in Jackson et al. (1994). Conditions (3) and (4) are slightly stronger than their assumptions (A1) and (A2) respectively, in that conditions (3) and (4) require two punishment levels instead of one; condition (5) is the same as assumption (A3). Separable environments are a special case of economic environments as defined by Jackson (1991); the environment considered in Section 5 of Moore and Repullo (1988) is separable.

An example of a separable environment is a public good provision problem with quasilinear utilities, where the mechanism designer can impose sufficiently large fines on the agents. Another example is a pure exchange economy with free disposal, where the utility of every agent is weakly increasing in every good and strictly increasing in at least one good. This environment is separable relative to a SCF f if there exists $\varepsilon > 0$ such that each agent gets at least ε of every good according to every allocation in $f(\Theta)$.

3. THE MAIN RESULT

In this section, we prove that in a separable environment, any monotonic SCF is Nash implementable with a mechanism that does not make use of integer or modulo games. Moreover, in any state $\theta \in \Theta$, any strategy that survives one round of elimination of weakly dominated

strategies and one round of elimination of strictly dominated strategies gives rise to the outcome equal to $f(\theta)$. This implies that the mechanism does not have any mixed-strategy Nash equilibria in undominated strategies that result in undesirable outcomes.

The main result of the note is

Proposition 1. *Suppose that environment \mathcal{E} is separable relative to a monotonic SCF f . Then f is implementable in Nash equilibrium by a mechanism that employs no integer or modulo games. In addition, for any $\theta \in \Theta$, any strategy profile that survives one round of elimination of weakly dominated strategies and one round of elimination of strongly dominated strategies results in the outcome $f(\theta)$.*

The outline of the construction is as follows. The agents are arranged on a circle and simultaneously make a report. A report of agent i may be either a pair of types, one for agent i and one for $i + 1$ (i 's right neighbor), or a pair of social alternatives. If all agents report pairs of types, then each agent i 's report of $i + 1$'s type is regarded as the truth. The mechanism is constructed so that telling the truth about one's type weakly dominates misrepresenting it or reporting an alternative. In addition, each agent is punished whenever he disagrees with his neighbor about the neighbor's type; therefore, if agent i telling the truth about his type, agent $i - 1$ has an incentive to tell the truth about i 's type as well.

What makes the mechanism different from the ones already in the literature (e.g. Jackson et al. (1994)) is that it does not have any pure-strategy Nash equilibria in weakly dominated strategies that lead to undesirable outcomes. In particular, there are no Nash equilibria where more than two agents make reports that don't match one another.

This is achieved by making use of the two punishment levels defined by assumptions (3) and (4). To understand how the mechanism uses differential punishments, suppose that agent 1 knows that all other agents agree with their neighbors, except agents 3 and 4, who disagree on agent 4's type. The mechanism has to be able to test whether it is agent 3 or agent 4 that is reporting the truth; at the same time, agent 1 has to be given an incentive to report his type truthfully. In a separable environment, this is possible to do with the help of punishments. In particular, in the situation described above, the mechanism makes agent 1's utility independent of his report of his own type: agent 4 gets a test allocation, and agent 1 is punished no matter what he reports. However, punishing an agent whenever other agents disagree may introduce undesirable Nash equilibria where all agents disagree with each other, knowing that they will be punished no matter what they report. To avoid this, the mechanism uses two punishment levels. The heavier punishment is invoked

whenever an agent disagrees with his neighbor on the neighbor's type; the lighter punishment is invoked whenever an agent agrees with his neighbor on the neighbor's type, but other agents disagree between each other. Therefore, each agent always has a strict incentive to match his neighbor's report, no matter what the other agents report.

The main drawback of the mechanism is that it employs large fines (although only off the equilibrium path). While it may be realistic in some applications that the social planner can severely punish the agents, it need not be so in others (for example, when the agents are liquidity constrained). Another difficulty that arises from the use of large fines is that the planner may not be able to commit to punish the agents. If the planner cannot commit to destroying resources ex post, agents will take that into account when making the reports, and that might change their incentives to report the truth.

Proof of Proposition 1. Let us say that $i \in I(\theta, \theta'_i)$ for some $i \in N, \theta \in \Theta, \theta'_i \in \Theta_i$ if $\exists b \in \mathcal{A}$ s.t. $u_i(\theta_i, f(\theta)) \geq u_i(\theta_i, b)$ and $u_i(\theta'_i, b) > u_i(\theta'_i, f(\theta))$. The monotonicity of f implies that $i \in I(\theta, \theta'_i)$ whenever $f(\theta) \neq f(\theta'_i, \theta_{-i})$, and that if $i \notin I(\theta, \theta'_i)$, then $f(\theta) = f(\theta'_i, \theta_{-i})$. For every $i \in N$, let B^i be the correspondence from $\Theta \times \Theta_i$ to \mathcal{A} such that $B^i(\theta, \theta'_i) = \{b \in \mathcal{A} : u_i(\theta_i, f(\theta)) \geq u_i(\theta_i, b) \text{ and } u_i(\theta'_i, b) > u_i(\theta'_i, f(\theta))\}$; if $i \in I(\theta, \theta'_i)$, then $B^i(\theta, \theta'_i)$ is nonempty. For every $i \in N$, let $b^i(\theta, \theta'_i)$ be a selection from the correspondence $B^i(\theta, \theta'_i)$ such that $b^i(\theta, \theta'_i)$ is preferred by type θ'_i to any other value of the selection which lies in $B^i(\theta, \theta'_i)$: i.e. for every $\theta, \tilde{\theta} \in \Theta, \theta'_i, \tilde{\theta}'_i \in \Theta_i$,

$$(6) \quad b^i(\tilde{\theta}, \tilde{\theta}'_i) \in B^i(\theta, \theta'_i) \Rightarrow u_i(\theta'_i, b^i(\theta, \theta'_i)) \geq u_i(\theta'_i, b^i(\tilde{\theta}, \tilde{\theta}'_i))$$

The fact that Θ is finite guarantees that such a selection $b^i(\theta, \theta'_i)$ exists.

Let $\mathcal{A}_+ = \{a \in \mathcal{A} : \forall i \in N, \theta_i \in \Theta_i, u_i(\theta_i, a) \geq u_i(\theta_i, \bar{w})\}$; \mathcal{A}_+ is nonempty, because $f(\Theta) \subseteq \mathcal{A}_+$. Consider the following mechanism $\Gamma = ((S_i)_{i \in N}, g)$. Let $S_i = \Theta_i \times \Theta_{i+1} \cup \mathcal{A}_+^2$, where $N + 1$ is taken to equal 1. Let $s_i \in \{(\alpha_i, \beta_i), (x_i, y_i)\}$ be a representative element of S_i , where $\alpha_i \in \Theta_i, \beta_i \in \Theta_{i+1}, x_i, y_i \in \mathcal{A}_+ : \alpha_i$ and β_i are agent i 's report on his own type and the type of agent $i + 1$, respectively, and (x_i, y_i) is a pair of social alternatives. Let $\beta = (\beta_i)_{i \in N}$, $\alpha = (\alpha_i)_{i \in N}$ and for any $j \in N$, let $\beta^{-j} = (\beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_N)$. For any report profile $s \in S = \times_{i \in N} S_i$, let $L = \{l \in N : s_l = (x_l, y_l)\}$ (the set of agents who report a pair of alternatives); $K = \{k \in N : s_k = (\alpha_k, \beta_k), s_{k+1} = (\alpha_{k+1}, \beta_{k+1}), \beta_k \neq \alpha_{k+1}\}$; and $J = \{j \in N : s_j = (\alpha_j, \beta_j), s_{j-1} = (\alpha_{j-1}, \beta_{j-1}), \beta_{j-1} \neq \alpha_j \text{ and } j \in I(\beta, \alpha_j)\}$. These definitions imply that if $j \in J$, then $j - 1 \in K$, and if $K = \emptyset$, then $J = \emptyset$; the converse of either statement need not hold.

The function $g(s)$ will be defined as follows:

I. If $L = \emptyset$, then:

I.1. If $J = \emptyset$, then $g(s) = [f(\beta)]^{N \setminus K, K}$ (i.e. $g(s)$ is an outcome that is equivalent to $f(\beta)$ for agents not in K and to \underline{w} for agents in K).

I.2. If $J = \{j\}$ for some $j \in N$, then:

I.2.a. If $j \notin K$, then $g(s) = [b^j(\beta, \alpha_j)]^{\{j\}, K}$ (i.e. $g(s)$ is an outcome that is equivalent to $b^j(\beta, \alpha_j)$ for agent j , to \underline{w} for agents in K and to \bar{w} for every other agent).

I.2.b. If $j \in K$, then $g(s) = [\bar{w}]^{N \setminus K, K}$ (i.e. $g(s)$ is an outcome that is equivalent to \bar{w} for agents not in K and to \underline{w} for agents in K).

I.3. If $|J| \geq 2$, then $g(s) = [\bar{w}]^{N \setminus K, K}$.

II. If $L \neq \emptyset$, then:

II.1. If $L = N \setminus \{i\}$, then $g(s) = \begin{cases} [x_{i-1}]^{\{i\}, N \setminus \{i\}}, & \text{if } u_i(\alpha_i, x_{i-1}) \geq u_i(\alpha_i, y_{i-1}); \\ [y_{i-1}]^{\{i\}, N \setminus \{i\}}, & \text{if } u_i(\alpha_i, y_{i-1}) > u_i(\alpha_i, x_{i-1}) \end{cases}$

II.2. Otherwise $g(s) = [\bar{w}]^{N \setminus \{K \cup L\}, K \cup L}$

We are going to show that:

(I) For any $\theta \in \Theta$, there exists a Nash equilibrium of this mechanism that results in the outcome $f(\theta)$.

(II) For any $\theta \in \Theta$, the mechanism has no pure-strategy Nash equilibria that result in an outcome different from $f(\theta)$.

(III) For any $\theta \in \Theta$, any strategy profile that survives one round of elimination of weakly dominated strategies and one round of elimination of strongly dominated strategies results in the outcome $f(\theta)$.

Proof of (I). Suppose that all agents but i play the truthful strategy, that is, $\forall j \neq i$, $(\alpha_j, \beta_j) = (\theta_j, \theta_{j+1})$. All strategies available to agent i can be classified into three groups:

(a) The truthful strategy $s_i = (\alpha_i, \beta_i) = (\theta_i, \theta_{i+1})$, which induces the outcome $f(\theta)$.

(b) Strategies $s_i = (\alpha_i, \beta_i)$ such that $\beta_i \neq \alpha_{i+1}$, or $s_i = (x_i, y_i)$. Then $i \in K \cup L$, and the outcome is equivalent to \underline{w} for him.

(c) Strategies such that $\alpha_i \neq \beta_{i-1} = \theta_i$ and $\beta_i = \alpha_{i+1} = \theta_{i+1}$. Then $i \notin K$ and $L = \emptyset$.

There are two cases. If $i \notin I(\beta, \alpha_i)$, then agent i 's utility equals $u_i(\theta_i, f(\beta))$, which is the same as what he would get if he announced $\alpha_i = \theta_i$. If $i \in I(\beta, \alpha_i)$, then $i \in J$, and i 's utility equals $u_i(\theta_i, b^i(\beta, \alpha_i)) = u_i(\theta_i, b^i(\theta, \alpha_i)) \leq u_i(\theta_i, f(\theta))$. So the deviations in this class are unprofitable as well.

Thus, $s_i = (\theta_i, \theta_{i+1})$ is a best response for i , and the resulting outcome is $f(\theta)$.

Proof of (II). Suppose that $s \in S$ is a Nash equilibrium of the mechanism. If for some $i \in N$, $\beta_i \neq \alpha_{i+1}$ or $s_i = (x_i, y_i)$, then $i \in K \cup L$ and the outcome is equivalent to \underline{w} for i , regardless of s_{-i} . A deviation to $s_i = (\alpha_i, \beta_i)$ such that $\alpha_i = \theta_i, \beta_i = \alpha_{i+1}$ results in a strictly better outcome for i , regardless of s_{-i} . Therefore for every $i \in N$, $\beta_i = \alpha_{i+1}$ in equilibrium, and the outcome is $f(\beta)$. If $f(\beta) \neq f(\theta)$, then, by the monotonicity of f , there exists an agent $i \in N$, for whom announcing $\alpha_i = \theta_i$ is a profitable deviation.

Proof of (III). Fix any agent $i \in N$. Let us prove that, for any $s_{-i} \in \times_{j \in N \setminus \{i\}} S_j$, and any $(\alpha_i, \beta_i) \in \Theta_i \times \Theta_{i+1}$,

$$(7) \quad u_i(\theta_i, g((\theta_i, \beta_i), s_{-i})) \geq u_i(\theta_i, g((\alpha_i, \beta_i), s_{-i})),$$

and for any $(x_i, y_i) \in \mathcal{A}_+^2, \beta_i \in \Theta_{i+1}$,

$$(8) \quad u_i(\theta_i, g((\theta_i, \beta_i), s_{-i})) \geq u_i(\theta_i, g((x_i, y_i), s_{-i}))$$

Moreover, for any $(\alpha_i, \beta_i) \in \Theta_i \times \Theta_{i+1}$ such that $\alpha_i \neq \theta_i$, there exists $s_{-i} \in \times_{j \in N \setminus \{i\}} S_j$ such that

$$(9) \quad u_i(\theta_i, g((\theta_i, \beta_i), s_{-i})) > u_i(\theta_i, g((\alpha_i, \beta_i), s_{-i}))$$

and for any $(x_i, y_i) \in \mathcal{A}_+^2, \beta_i \in \Theta_{i+1}$, there exists $s_{-i} \in \times_{j \in N \setminus \{i\}} S_j$ such that

$$(10) \quad u_i(\theta_i, g((\theta_i, \beta_i), s_{-i})) > u_i(\theta_i, g((x_i, y_i), s_{-i}))$$

These inequalities imply that all strategies $s_i \in \mathcal{A}_+^2$ and all strategies $s_i = (\alpha_i, \beta_i) \in \Theta_i \times \Theta_{i+1}$ such that $\alpha_i \neq \theta_i$ are weakly dominated.

First, note that any $s_i = (x_i, y_i)$ leads to an outcome equivalent to \underline{w} for i , regardless of s_{-i} , whereas $s_i = (\theta_i, \beta_i)$ for any $\beta_i \in \Theta_{i+1}$ always leads to an outcome at least as good as \underline{w} . This proves inequality (8). Fix any $(x_i, y_i) \in \mathcal{A}_+^2, \beta_i \in \Theta_{i+1}$. If $\alpha_{i+1} = \beta_i$ and $\forall j \in N \setminus \{i, i+1\}, \beta_{j-1} = \alpha_j$, then announcing (x_i, y_i) leads to an outcome equivalent to \underline{w} for i , whereas announcing (θ_i, β_i) leads to an outcome strictly better than \underline{w} . This proves inequality (10).

To prove inequalities (7) and (9), fix any $(\alpha_i, \beta_i) \in \Theta_i \times \Theta_{i+1}$.

Case I. Suppose $\exists j \neq i : s_j = (x_j, y_j) \in \mathcal{A}_+^2$. There are three subcases:

1) $\forall j \neq i, s_j = (x_j, y_j) \in \mathcal{A}_+^2$. Then announcing (θ_i, β_i) leads to an outcome equivalent to x_{i-1} if $u_i(\theta_i, x_{i-1}) \geq u_i(\theta_i, y_{i-1})$ and equivalent to y_{i-1} otherwise. Therefore, announcing (θ_i, β_i) is at least as good as (α_i, β_i) . Moreover, by assumption (5), for every $\alpha_i \in \Theta_i$ there exists $(x_{i-1}, y_{i-1}) \in \mathcal{A}_+^2$ such that announcing (θ_i, β_i) is strictly better than (α_i, β_i) . The last point proves inequality (9).

2) Suppose $\exists k \neq i : s_k = (\alpha_k, \beta_k) \in \Theta_k \times \Theta_{k+1}$ and $s_{i+1} = (x_{i+1}, y_{i+1}) \in \mathcal{A}_+^2$. Then the outcome is equivalent to \bar{w} for i if $s_i = (\alpha_i, \beta_i)$ for any $(\alpha_i, \beta_i) \in \Theta_i \times \Theta_{i+1}$ and equivalent to \underline{w} otherwise. Therefore, announcing (θ_i, β_i) is equivalent to announcing any (α_i, β_i) .

3) Suppose $\exists k \neq i : s_k = (\alpha_k, \beta_k) \in \Theta_k \times \Theta_{k+1}$ and $s_{i+1} = (\alpha_{i+1}, \beta_{i+1})$. Then the outcome is equivalent to \bar{w} for i if $s_i = (\alpha_i, \beta_i)$ such that $\beta_i = \alpha_{i+1}$ and equivalent to \underline{w} otherwise. Therefore, announcing (θ_i, β_i) is equivalent to announcing (α_i, β_i) , for any $\alpha_i \in \Theta_i, \beta_i \in \Theta_{i+1}$.

Case II. Suppose that $\forall j \neq i, s_j = (\alpha_j, \beta_j) \in \Theta_j \times \Theta_{j+1}$. There are four subcases:

1) Suppose that $\forall j \neq i, \beta_{j-1} = \alpha_j$ (in particular, $\beta_i = \alpha_{i+1}$) and $\beta_{i-1} = \theta_i$. Then announcing $\alpha_i = \theta_i$ leads to the outcome $f(\beta)$; announcing $\alpha_i \neq \theta_i$ leads to either an outcome equivalent to $f(\beta)$ for i (if $i \notin I(\beta, \alpha_i)$) or an outcome equivalent to $b^i(\beta, \alpha_i)$ for i (if $i \in I(\beta, \alpha_i)$). But $\beta_{i-1} = \theta_i$, so, if $i \in I(\beta, \alpha_i)$, then $u_i(\theta_i, b^i(\beta, \alpha_i)) \leq u_i(\theta_i, f(\beta))$ by the definition of b^i .

2) Suppose that $\forall j \neq i, \beta_{j-1} = \alpha_j$ (in particular, $\beta_i = \alpha_{i+1}$, so that $i \notin K$) and $\beta_{i-1} \neq \theta_i$. Then announcing $\alpha_i = \theta_i$ gives i either $u_i(\theta_i, f(\beta))$ (if $i \notin I(\beta, \theta_i)$) or $u_i(\theta_i, b^i(\beta, \theta_i))$ (if $i \in I(\beta, \theta_i)$). Announcing $\alpha_i = \beta_{i-1} \neq \theta_i$ gives i the utility of $u_i(\theta_i, f(\beta)) < u_i(\theta_i, b^i(\beta, \theta_i))$ by the definition of $b^i(\beta, \theta_i)$, so it cannot be more profitable than announcing $\alpha_i = \theta_i$. Announcing $\alpha_i \notin \{\beta_{i-1}, \theta_i\}$ gives i the utility of either $u_i(\theta_i, f(\beta))$ (if $i \notin I(\beta, \alpha_i)$) or $u_i(\theta_i, b^i(\beta, \alpha_i))$ (if $i \in I(\beta, \alpha_i)$). There are four possibilities:

(a) $i \notin I(\beta, \theta_i), i \notin I(\beta, \alpha_i)$.

Then announcing θ_i gives i the same utility as announcing α_i .

(b) $i \notin I(\beta, \theta_i), i \in I(\beta, \alpha_i)$.

Then, by definition of $I(\beta, \theta_i)$,

$$\nexists b \in \mathcal{A} : \begin{cases} u_i(\beta_{i-1}, f(\beta)) \geq u_i(\beta_{i-1}, b), \\ u_i(\theta_i, b) > u_i(\theta_i, f(\beta)). \end{cases}$$

This, together with the fact that

$$u_i(\beta_{i-1}, f(\beta)) \geq u_i(\beta_{i-1}, b^i(\beta, \alpha_i)),$$

implies that

$$u_i(\theta_i, f(\beta)) \geq u_i(\theta_i, b^i(\beta, \alpha_i))$$

so announcing θ_i is weakly better for i than announcing α_i .

(c) $i \in I(\beta, \theta_i), i \notin I(\beta, \alpha_i)$.

Then announcing θ_i is strictly better for i than announcing α_i by the definition of $b^i(\beta, \theta_i)$.

(d) $i \in I(\beta, \theta_i)$, $i \in I(\beta, \alpha_i)$. We will prove that

$$u_i(\theta_i, b^i(\beta, \theta_i)) \geq u_i(\theta_i, b^i(\beta, \alpha_i)),$$

so that announcing θ_i is at least as good for i as announcing α_i . Suppose not; then

$$u_i(\theta_i, b^i(\beta, \alpha_i)) > u_i(\theta_i, b^i(\beta, \theta_i)) > u_i(\theta_i, f(\beta)),$$

where the last inequality follows from the definition of $b^i(\beta, \theta_i)$. Also,

$$u_i(\beta_{i-1}, f(\beta)) \geq u_i(\beta_{i-1}, b^i(\beta, \alpha_i))$$

by the definition of $b^i(\beta, \alpha_i)$. It follows that

$$b^i(\beta, \alpha_i) \in B^i(\beta, \theta_i)$$

But then property (6) of the selection b^i implies

$$u_i(\theta_i, b^i(\beta, \theta_i)) \geq u_i(\theta_i, b^i(\beta, \alpha_i))$$

– contradiction.

3) Suppose that $\exists j \neq i : \beta_{j-1} \neq \alpha_j$ and $\forall k \notin \{i, j\}, \beta_{k-1} = \alpha_k$. Then there are two cases:

(a) Suppose that $j = i + 1$. Then $i \in K$, so i gets an outcome equivalent to \underline{w} no matter what α_i he announces.

(b) Suppose that $j \neq i + 1$ (so that $i \notin K$). If $j \in I(\beta, \alpha_j)$, then i gets an outcome equivalent to \bar{w} no matter what α_i he announces. If $j \notin I(\beta, \alpha_j)$, then this case is identical to case 1) if $\beta_{i-1} = \theta_i$, or case 2) if $\beta_{i-1} \neq \theta_i$.

4) Suppose that $|j \in N \setminus \{i\} : \beta_{j-1} \neq \alpha_j| \geq 2$. Then there are two cases:

(a) Suppose that $\beta_i \neq \alpha_{i+1}$. Then i gets an outcome equivalent to \underline{w} no matter what α_i he announces.

(b) Suppose that $\beta_i = \alpha_{i+1}$. If $\{j \in N \setminus \{i\} : \beta_{j-1} \neq \alpha_j \text{ and } j \in I(\beta, \alpha_j)\} = \emptyset$, then this case is identical to case 1) if $\beta_{i-1} = \theta_i$, or case 2) if $\beta_{i-1} \neq \theta_i$. If

$|\{j \in N \setminus \{i\} : \beta_{j-1} \neq \alpha_j \text{ and } j \in I(\beta, \alpha_j)\}| = 1$, then this case is identical to case 3(b). If $|\{j \in N \setminus \{i\} : \beta_{j-1} \neq \alpha_j \text{ and } j \in I(\beta, \alpha_j)\}| \geq 2$, then i gets an outcome equivalent to \bar{w} no matter what α_i he announces.

This completes the proof of inequalities (7) – (10), which imply that all strategies s_i such that either $s_i \in \mathcal{A}_+^2$ or $s_i = (\alpha_i, \beta_i) \in \Theta_i \times \Theta_{i+1}$ and $\alpha_i \neq \theta_i$ are weakly dominated.

Conversely, no strategy $s_i = (\alpha_i, \beta_i) \in \Theta_i \times \Theta_{i+1}$ such that $\alpha_i = \theta_i$ is weakly dominated. Suppose that $s_i = (\theta_i, \beta_i)$ and $s'_i = (\theta_i, \beta'_i)$ for $\beta_i, \beta'_i \in \Theta_{i+1}$. Let $s_{-i} = (\alpha_j, \beta_j)_{j \neq i}$ be

such that $\forall j \in N \setminus \{i, i+1\}, \beta_{j-1} = \alpha_j$ and $\alpha_{i+1} = \beta_i$ and $s'_{-i} = (\alpha'_j, \beta'_j)_{j \neq i}$ be such that $\forall j \in N \setminus \{i, i+1\}, \beta'_{j-1} = \alpha'_j$ and $\alpha'_{i+1} = \beta'_i$. Then $u_i(\theta_i, g(s_i, s_{-i})) > u_i(\theta_i, g(s'_i, s_{-i}))$ and $u_i(\theta_i, g(s'_i, s'_{-i})) > u_i(\theta_i, g(s_i, s'_{-i}))$. Therefore, neither s_i is weakly dominated by s'_i , nor vice versa.

Now suppose that every $i \in N$ always plays a strategy (α_i, β_i) such that $\alpha_i = \theta_i$. Then for every $i \in N$, any strategy (α_i, β_i) such that $\beta_i \neq \theta_{i+1}$ is strongly dominated, as it results in an outcome that is equivalent to \underline{w} for i , whereas $(\alpha_i, \beta_i) = (\theta_i, \theta_{i+1})$ results in an outcome strictly better than \underline{w} , no matter what β_{-i} is. Therefore, the only strategy that survives is $s_i = (\theta_i, \theta_{i+1})$. ■

REFERENCES

- [1] Abreu, D., Matsushima, H., 1992a. Virtual implementation in iteratively undominated strategies: complete information. *Econometrica* 60, 993-1008.
- [2] Abreu, D., Matsushima, H., 1992b. Virtual implementation in iteratively undominated strategies: incomplete information. Mimeo.
- [3] Abreu, D., Matsushima, H., 1994. Exact implementation. *Journal of Economic Theory* 64, 1-19.
- [4] Arya, A., Glover, J., Young, R., 1995. Virtual implementation in separable Bayesian environments using simple mechanisms. *Games and Economic Behavior* 9., 127-138.
- [5] Börgers, T., 1994, Weak dominance and approximate common knowledge. *Journal of Economic Theory* 64, 265-276.
- [6] Dekel, E., Fudenberg, D., 1990. Rational behavior with payoff uncertainty. *Journal of Economic Theory* 52, 243-267.
- [7] Duggan, J., 2003. Nash implementation with a private good. *Economic Theory* 21, 117-131.
- [8] Duggan, J., Roberts, J., 2002. Implementing the efficient allocation of pollution. *American Economic Review* 92, 1070-1078.
- [9] Jackson, M.O., 1991. Bayesian implementation. *Econometrica* 59, 461-477.
- [10] Jackson, M.O., 1992. Implementation in undominated strategies: a look at bounded mechanisms. *Review of Economic Studies* 59, 757-775.
- [11] Jackson, M.O., Palfrey, T., Srivastava, S., 1994. Undominated Nash implementation in bounded mechanisms. *Games and Economic Behavior* 6, 474-501.
- [12] Maskin, E., 1999. Nash equilibrium and welfare optimality. *Review of Economic Studies* 66, 23-38.
- [13] Mezzetti, C., Renou, L., 2010. Implementation in Mixed Nash Equilibrium. Mimeo.
- [14] Mizukami, H., Wakayama, T., 2007. Dominant strategy implementation in economic environments. *Games and Economic Behavior* 60, 307-325.
- [15] Moore, J., Repullo, R., 1988. Subgame perfect implementation. *Econometrica* 56, 1191-1220.
- [16] Saijo, T., 1988. Strategy space reduction in Maskin's theorem: Sufficient conditions for Nash implementation. *Econometrica* 56, 693-700.
- [17] Saijo, T., Sjöström, T., Yamato, T., 2007. Secure implementation. *Theoretical Economics* 2, 203-229.

- [18] Serrano, R., Vohra, R., 2010. Multiplicity of mixed equilibria in mechanisms: A unified approach to exact and approximate implementation. *Journal of Mathematical Economics* 46, 775-785.
- [19] Sjöström, T., 1994. Implementation in undominated Nash equilibria without integer games. *Games and Economic Behavior* 6, 502-511.
- [20] Sjöström, T., 1996. Credibility and renegotiation of outcome functions in implementation. *Japanese Economic Review* 47, 157-169.
- [21] Sjöström, T., 1999. Undominated Nash implementation with collusion and renegotiation. *Games and Economic Behavior* 26, 337-352.
- [22] Yamato, T., 1993. Double implementation in Nash and undominated Nash equilibria. *Journal of Economic Theory* 59, 311-323.

DEPARTMENT OF ECONOMICS, UNIVERSITY OF WESTERN ONTARIO, SOCIAL SCIENCE CENTRE, LONDON,
ONTARIO N6A 5C2, CANADA; MGOLTSMA@UWO.CA