

Available online at www.sciencedirect.com



JOURNAL OF Economic Theory

Journal of Economic Theory 133 (2007) 111-126

www.elsevier.com/locate/jet

Subgame perfect implementation: A full characterization

Hannu Vartiainen*

Yrjö Jahnsson Foundation, Ludviginkatu 3-5, 00130 Helsinki, Finland

Received 30 January 2001; final version received 27 June 2005 Available online 14 November 2005

Abstract

Moore and Repullo [Subgame perfect implementation, Econometrica 56 (1988) 1191–1220], and Abreu and Sen [Subgame perfect implementation: a necessary and almost sufficient condition, J. Econ. Theory 50 (1990) 285–299] introduce distinct necessary and sufficient conditions for SPE implementation, when the number of players is at least three. This paper closes the gap between the conditions—a complete characterization of the SPE implementable choice rules is provided. The characterization consists of α^* , which strengthens α of Abreu–Sen by adding it a restricted veto-power condition, and the *unanimity* condition. Under strict preferences α^* is equal to α .

© 2005 Elsevier Inc. All rights reserved.

JEL classification: D82; D72

Keywords: Subgame perfect equilibrium; Implementation; Characterization

1. Introduction

Moore and Repullo [14] (hereafter MR) and Abreu and Sen [2] (hereafter AS) study implementation of social choice rules (correspondences) by using *multi-stage mechanisms* and *subgame perfect Nash equilibrium* (SPE) as the solution concept.¹ They demonstrate that having multiple stages in the implementing mechanism is very helpful: much more can be subgame perfect implemented than Nash implemented.²

0022-0531/\$ - see front matter 0 2005 Elsevier Inc. All rights reserved. doi:10.1016/j.jet.2005.06.008

^{*} Fax: +358 9 605002.

E-mail address: hannu.vartiainen@yjs.fi.

¹ For recent introduction, see [13] or [11].

 $^{^{2}}$ Tree mechanisms, where backwards induction can be used, are a special case of multi-stage mechanisms. Moulin [16] is a seminal treatment on backwards induction. Herrero and Srivastava [9] give a full characterization.

AS and MR characterize general conditions for SPE implementation in the three-or-more players case. AS identifies *Condition* α that bears the same relation to SPE implementation as Maskin's [12] *monotonicity* condition to Nash implementation. I.e. α is a *necessary* condition and coupled with *no-veto power* (NVP) a *sufficient* condition for SPE implementation.³

In this paper, we fill the gap between AS necessary and sufficient conditions.⁴ Our characterization consists of a strengthened version of α , which we dub *Condition* α^* , and the *unanimity* condition. It is shown that if a choice rule is SPE implementable, then it satisfies the latter two properties. To prove the other direction, we employ a modification of the stage-mechanism introduced by MR. We show that if a choice rule meets α^* and unanimity, then our MR-mechanism SPE implements it.⁵ Thus, this mechanism is *canonical*.

To understand the content of Condition α^* , suppose that f can be SPE implemented. Suppose further that a belongs to $f(\phi)$ but not in $f(\theta)$. Then there is an equilibrium of an implementing mechanism that induces a under ϕ but not under θ . Condition α^* identifies situations when a profitable deviation under θ from the equilibrium path under ϕ is *not* followed by further deviations by a player different from the originally deviating player, say i. In such case we say i's deviations are *successively unilateral*. If we know that player i's θ -maximal successively unilateral sequence of deviations leads to the implementation of outcome b which is also all j's, $j \neq i$, θ -maximal in the range of the implementing mechanism, then b must be an equilibrium outcome under θ . If f is SPE implemented by the mechanism, then we must have $b \in f(\theta)$. Thus α^* restricts the feasible outcomes when only successively unilateral deviations are possible, and when deviator's maximal deviation sequences lead to the implementation of an outcome that is maximal also to all non-deviators.

Clearly α^* is stronger than α : It restricts choice rules when α does but also in the subset of those cases where n - 1 players agree on a maximal alternative. Thus α^* can be understood as a combination of α and a *restricted veto-power* condition. However, α^* is strictly weaker than α combined with NVP.

The weakness of the extra restriction imposed by α^* relative to α becomes transparent when we assume *linear* preferences. It is shown that the restricted veto-power part of α^* is *vacuously* satisfied whenever α is, i.e. α^* *coincides* with α . This means that α and unanimity are, in fact, a *necessary and sufficient* condition for SPE implementation under linear preferences. We also show that α simplifies further when preferences are linear.

The paper is organized as follows: Section 2 introduces the basic concepts and the equilibrium notion we use. Section 3 provides the main characterization result and establishes some simplifications and corollaries. Section 4 concludes the paper.

³ Srivastava and Trick [20] characterize rules that are implementable via multi-stage mechanisms where a stage is a binary voting contest. Under sophisticated, i.e. sincere given perfect foresight, voting, players' preferences can be summarized by a tournament. This permits Srivastava and Trick to give an elegant characterization of implementable rules. Their result is not implied nor does it imply the results of this paper. The former follows from the *assumption* that players vote sincerely given their perfect foresight, which is a refinement on the equilibrium concept. The latter is due to the fact that the underlying game in Srivastava–Trick is not a canonical multi-stage game.

⁴ In the context of Nash implementation, Moore and Repullo [15] fill the gap between Maskin's necessary and sufficient conditions.

⁵ The mechanism uses integer game (see e.g. [10,11] for discussion). The usual counterargument to the criticism is that general results require abstract mechanisms. Recent papers using integer games include Bergin and Sen [4], and Brusco [5,6].

2. Fundamentals

2.1. Notation and the basic set up

Let *I* be the set of players, with generic elements *i*, *j*. Assume $\#I \ge 3$. There is set *A* of feasible pure alternatives with $\#A \ge 3$ and set Θ of admissible *preferences* of players. For profile $\theta \in \Theta$, player *i*'s preferences over *A* are represented by $R_i(\theta)$ ($P_i(\theta)$ represents a strict relation).⁶ Each θ is a complete description of state, and θ is observable by all players but not by others. Denote the *lower contour set* of *i* at $a \in A$ under θ by $L_i(a, \theta) = \{b \in A : aR_i(\theta)b\}$. Denote by $M_i(\theta, B)$ the set of *i*'s θ -maximal alternatives in $B \subseteq A$ under θ . Formally, $M_i(\theta, B) = \{a \in B : aR_i(\theta)b\}$ for all $b \in B\}$.

We say that *i* experiences a *preference reversal between a and b* when switching from ϕ to θ if $aR_i(\phi)b$ and $bP_i(\theta)a$. Thus, if *i* experiences a preference reversal between *a* and some element in *B* when switching from ϕ to θ , then $[L_i(a, \phi) \setminus L_i(a, \theta)] \cap B \neq \emptyset$. Preference reversal is a fundamental requirement for implementation.

2.2. SPE Implementation and a choice rule

We confine our attention to mechanisms which are representable by *extensive game forms* with simultaneous moves.⁷ A mechanism is an array $\Gamma = \langle Y, S, g \rangle$ where Y is a set of histories $y, S = S_1 \times \cdots \times S_n$ and $S_i = \times_{y \in Y} S_i^y$ for all i. An element of $S^y = S_1^y \times \cdots \times S_n^y$, say $s^y = (s_1^y, \ldots, s_n^y)$, is a message vector while s_i^y is i's message at y. Histories and messages are tied together by the property that $S^y = \{s^y : (y, s^y) \in Y\}$. An element of S_i , say s_i , is i's (pure) strategy,⁸ specifying i's choices at each non-terminal history, and an element of S, say s, is a (pure) strategy profile. There is an initial history $\emptyset \in Y$ and each history $y^k \in Y$ is represented by a finite sequence $(\emptyset, s^1, \ldots, s^{k-1}) = y^k$.⁹ If $y^{k+1} = (y^k, s^k)$, then history y^{k+1} proceeds history y^k . As Γ contains finitely many stages, there is a set of terminal histories \overline{Y} of Y such that $\overline{Y} = \{y \in Y$: there is no $y' \in Y$ proceeding y}. Any strategy profile $s \in S$ defines a unique terminal history given the initial history y. Sometimes we call terminal history a path.

The *outcome function* $g: S \to A$ specifies an outcome for each terminal history, and hence for each strategy profile. Given state $\theta \in \Theta$, the pair (Γ, θ) constitutes an extensive form game with simultaneous moves.

By the construction of Γ , every $y \in Y \setminus \overline{Y}$ identifies a *subgame* $\Gamma(y)$ of Γ , as follows: y is an initial history of the game $\Gamma(y) = \langle Y(y), S(y), g \rangle$ where $Y(y) = \{y' \in Y : y' \text{ proceeds } y\}$ and $S(y) = \times_{y' \in Y(y)} S^{y'}$. Then $S(y) = S_1(y) \times \cdots \times S_n(y)$, where $S_i(y)$ is the strategy set of i in subgame $\Gamma(y)$. Denote an element of S(y) by s : y. Now s : y specifies a unique terminal history. The corresponding value of the outcome function is written g(s : y). Thus $g(s) = g(s : \emptyset)$. Hence $\Gamma = \Gamma(y^0)$. Denote by $D(y, s_{-i})$ the set of outcomes player i can reach by varying his own

⁶ A preference relation is a complete and transitive binary relation.

⁷ For an equivalent definition, see [17, Chapter 6].

⁸ For simplicity, we confine our attention to pure strategies. This restriction does *not* affect the results.

⁹ Thus, like MR, we confine our attention to games having *finitely* many stages. This assumption is made for sake of simplicity. AS showed that allowing infinitely many stages should not affect the conclusions.

strategy $s_i \in S_i$ given that history y is reached, and all $j \neq i$ adopt strategy $s_i \in S_j$. Formally,

$$D(y, s_{-i}) = \{a \in A : g(s : y) = a, \text{ for some } s_i \in S_i(y)\}.$$

Denote by $\overline{SPE}(\Gamma, \theta)$ the set of SPE *strategies* of a game Γ given θ .

Definition 1. Take $\Gamma = \langle Y, S, g \rangle$ and $\theta \in \Theta$. Then $s \in S$ is an element of $\overline{SPE}(\Gamma, \theta)$ if and only if $g(s : y) \in M_i(\theta, D(y, s_{-i}))$ for all $y \in Y$ and for all $i \in I$.

By construction, $\overline{SPE}(\Gamma(y), \theta)$ refers to the equilibria of the subgame starting at history y. An important and natural additional restriction on any applicable mechanism $\Gamma = \langle Y, S, g \rangle$ is that $\overline{SPE}(\Gamma(y), \theta)$ is non-empty for all $y \in Y$ and for all $\theta \in \Theta$.¹⁰ Thus a mechanism must be well defined independently of the history of the play.

Denote by $SPE(\Gamma, \theta)$ the set of terminal outcomes, arising from equilibrium strategies.

Definition 2. Take $\Gamma = \langle Y, S, g \rangle$ and $\theta \in \Theta$. Then $a \in A$ is an element of $SPE(\Gamma, \theta)$ if and only if g(s) = a for some $s \in \overline{SPE}(\Gamma, \theta)$.

Note that $SPE(\Gamma(y), \theta)$ refers to outcomes induced by SPE strategies after history *y* is reached. Of course, $SPE(\Gamma(y), \theta)$ and $SPE(\Gamma, \theta)$ do not typically coincide if *y* does not belong to the *equilibrium path*.

A social choice rule $f : \Theta \to A$ defines a non-empty set of socially "desirable" outcomes in each state. Denote $f(\Theta) = \bigcup_{\theta \in \Theta} f(\theta)$. We are interested in *full* implementation where the set of SPE outcomes of the implementing mechanism coincides with the desired choice rule in all states.

Definition 3. Choice rule *f* is SPE implemented by a mechanism Γ if and only if $SPE(\Gamma, \theta) = f(\theta)$ for all $\theta \in \Theta$.

We say that if for some choice rule f there is a Γ such that f is SPE implemented by Γ , then f is SPE *implementable*. If f is implementable, then there is $D \supseteq f(\Theta)$ such that D is the image of the implementing mechanism, that is D = g(S), where S, g are the components of Γ which SPE implements f. In what follows, D can always be interpreted to be the range of the implementing mechanism.

We want to characterize SPE implementable choice rules, where characterization means "without reference to the definition of SPE". That is, a characterization imposes a restriction on admissible rules as a direct function of basic data.

3. The characterization

First we review some fundamental properties of *Nash* implementable choice rules.¹¹ Choice rule *f* satisfies *monotonicity* with respect to $D \supseteq f(\Theta)$ if there *i* such that $[L_i(a, \phi) \setminus L_i(a, \theta)] \cap D \neq \emptyset$ whenever $a \in f(\phi) \setminus f(\theta)$, for any $\phi, \theta \in \Theta$, and it satisfies NVP with respect to $D \supseteq f(\Theta)$ if $b \in \cap_{i \neq i} M^j(\theta, D)$ for any $i \in I$ implies $b \in f(\theta)$. In many environments, including the

114

¹⁰ This restriction is weaker than the *best response property*, see e.g. Jackson [11].

¹¹ If we would restrict the mechanisms having a *single* stage, then SPE implementation would collapse back to Nash implementation.

"economic" ones, NVP is vacuously satisfied. By Maskin [12] we have the following. If choice rule f can be Nash implemented, then it satisfies monotonicity w.r.t. some $D \supseteq f(\Theta)$. If f satisfies monotonicity and NVP w.r.t. some $D \supseteq f(\Theta)$, then it can be Nash implemented.¹²

To express the AS characterization of SPE implementable choice rules, define the following pair of sequences. Given a triple $(a, \phi, \theta) \in A \times \Theta \times \Theta$ and $D \subseteq A$, a pair of sequences $h(0), \ldots, h(K)$ in I, and $d(0), \ldots, d(K+1)$ in D satisfy a = d(0) and

 $\alpha(i) \ d(k)R_{h(k)}(\phi)d(k+1), \text{ for } k = 0, \dots, K,$

 $\begin{aligned} &\alpha(\text{ii}) \ d(K+1)P_{h(K)}(\theta)d(K), \\ &\alpha(\text{iii}) \ d(k) \notin M^{h(k)}(\theta, D), \text{ for } k = 0, \dots, K, \end{aligned}$

 $\alpha(\text{iv}) \ d(K+1) \in \bigcap_{i \neq h(K)} M^i(\theta, D)$ implies that K = 0 or $h(K-1) \neq h(K)$.

Given $(a, \phi, \theta : D)$, denote a typical sequence meeting α -conditions by $(h, d)(a, \phi, \theta : D)$ (when there is no risk of confusion, we may drop the $(a, \phi, \theta : D)$ -part).

Definition 4 (*Conditions* α [α^{-}]). There is $D \supseteq f(\Theta)$ and, for all (a, ϕ, θ) such that $a \in f(\phi) \setminus$ $f(\theta)$, there is a sequence $(h, d)(a, \phi, \theta : D)$ meeting $\alpha(i)$ -(iv) $[\alpha(i)$ -(iii), resp.].

Parts $\alpha(i)$ and $\alpha(ii)$ were derived by MR, AS introduced $\alpha(iii)$ and $\alpha(iv)$. Note that $\alpha(i)$ and $\alpha(ii)$ imply that there is player h(K) who experiences a preference reversal between d(K) and d(K + 1), when switching from ϕ to θ . A need for preference reversal is acknowledged already by monotonicity. However, monotonicity requires that there is a preference reversal between a choice *in f* and some other alternative.

AS showed that α , and thus α^- , is a necessary condition for SPE implementation. The intuitive reason is that mechanism Γ that implements f gives rise to sequence $(h, d)(a, \phi, \theta : D)$ meeting $\alpha(i)$ -(iv). To see how, suppose that $d(0) = a \in f(\phi) \setminus f(\theta)$. Then there is player h(0), who deviates under θ from the path implementing d(0) under ϕ . Since the deviation must be strictly profitable, d(0) cannot be top ranked for h(0) in the range of the mechanism (part $\alpha(iii)$). On the other hand, there must be a ϕ -SPE outcome following the deviation, say d(1), that is not more profitable under ϕ for the deviator than d(0) (part $\alpha(i)$). Otherwise d(0) would not constitute a SPE at the first place. Ask if d(1) constitutes a θ -SPE at the subgame following the deviation. If yes, then d(1) must be more profitable to h(0) than d(0) under θ (part $\alpha(ii)$). If not, then d(1) constitutes a ϕ -SPE but not θ -SPE at the subgame following the deviation, and hence we can start the process from the beginning. Continuing this way, and noting that the process must end at some point, we come up with a sequence meeting $\alpha(i)$ -(iii). Thus, any deviation under θ from the path implementing a under ϕ , which is necessitated by $a \in f(\phi) \setminus f(\theta)$, induces a $(h, d)(a, \phi, \theta; D)$ meeting $\alpha(i)$ -(iii). Condition $\alpha(iv)$ follows from observation that if no sequence (h, d) induced by Γ meets $\alpha(iv)$, then one reduce the length of a some (h', d') by pruning all (h', d')(k), k at or below K such that h(k) = i for some i, to obtain sequence meeting all $\alpha(i)$ -(iv).

In addition, AS derive a sufficient condition for SPE implementation: if f satisfies Condition α and NVP, then it *can* be SPE implemented. We will show that when preferences are linear, this result holds with Condition α^{-} .

¹² The gap between Maskin's necessary and sufficient conditions was filled by Moore and Repullo [15] (see also [18,7]) by introducing Condition μ which is *both* necessary and sufficient for Nash implementation. In addition to monotonicity, Condition μ contains a *restricted* version of NVP, and a *unanimity* property. Our conditions below can be thought as generalization of these conditions.

Theorem 1 (Abreu and Sen [2]). If f can be SPE implemented, then it satisfies Condition α w.r.t. some $D \supseteq f(\Theta)$. Moreover, if $n \ge 3$, and f satisfies Condition α and NVP w.r.t. some $D \supseteq f(\Theta)$, then it can be SPE implemented.

This result is of particular interest by its similarity to Maskin's Theorem—Condition α plays the role of Maskin monotonicity in both sides of the characterization. Note that if one restricts the focus on (h, d) sequences of length K = 0, then Condition α reduces to Maskin monotonicity, and Theorem 1 back to Maskin's Theorem.

As there is a gap between the necessary and sufficient conditions, Theorem 1 leaves open the *full* characterization of SPE implementation. Our task is to fill the gap. To do that, we introduce conditions which are weaker than NVP, but still necessary for implementation.

Definition 5 (Unanimity). There is set D such that $D \supseteq f(\Theta)$, and such that if $b \in \bigcap_{i \in I} M^i(\theta, D)$, then $b \in f(\theta)$.

Of course, unanimity is strictly weaker than NVP.¹³ However, it is a necessary condition for SPE implementation: if there is an outcome which is unanimously top ranked in the range of the mechanism under certain preference profile, then it must be an equilibrium outcome under that profile because there is no player who would strictly benefit from deviation from the strategy implementing this outcome (by construction, such strategy exists).

Lemma 1. Choice rule f is SPE implementable only if it satisfies unanimity w.r.t. some D.

Proof. Let mechanism $\Gamma = \langle Y, S, g \rangle$ SPE implement f. Define D = g(S). In particular $f(\theta) \in D$ for all $\theta \in \Theta$. If $b \in \bigcap_{i \in I} M^i(\theta, D)$, then there is $s \in S$ such that b = g(s). Let $\bar{y}(s) = (y^0, s^1, \ldots, s^K)$ for some $K \ge 0$ and define $Y^K = \{y^k : y^k = (y^0, s^1, \ldots, s^k), k = 0, \ldots, K\}$. By assumption, one is able to find s' such that $\bar{y}(s) = \bar{y}(s')$, and such that $g(s' : y) \in SPE(\Gamma(y), \theta)$ for all $y \in Y \setminus Y^K$. Necessarily $b = g(s') \in \bigcap_{i \in I} D(y^k, s'_{-i})$ for all $k = 0, \ldots, K$, and since $D(y^k, s'_{-i}) \subseteq D$, for all $k = 0, \ldots, K, i \in I$, it follows that also $g(s') \in \bigcap_{i \in I} M_i(\theta, D(y^k, s_{-i}))$ for all $k = 0, \ldots, K$. Hence $s' \in \overline{SPE}(\Gamma, \theta)$ and, therefore, $b \in f(\theta)$. \Box

More restrictions must be imposed on choice rule to fill the gap between necessary and sufficient conditions. The condition we next introduce concerns circumstances under which d(K + 1) is top ranked by all but player h(K). It allows players to have "restricted veto power". The next example serves as a motivation.

Example 1. Let $I = \{1, 2, 3\}, \Theta = \{\phi, \theta\}, A = \{a, b, c, d, e\}$ and preference rankings are

$$R_1(\phi) = R_1(\theta) \quad R_2(\phi) = R_2(\theta) \quad R_3(\phi) \quad R_3(\theta)$$

a, b	a, b	d	С
С	С	е	d
d	d	С	е
е	е	а	b
		b	a

¹³ Unanimity is implied by $\mu(iii)$ of Moore and Repullo [15] on Nash implementable choice rules.

The Borda rule f^{B} chooses $f^{B}(\phi) = \{a\}$ and $f^{B}(\theta) = \{c\}$. Note that such rule satisfies Condition α (even monotonicity) and unanimity (but not NVP). Is f^{B} SPE implementable? Suppose yes. Then there is Γ which implements f^{B} . Since $a \in f^{B}(\phi) \setminus f^{B}(\theta)$, Γ contains a strategy which implements a, is a SPE under ϕ , and is not a SPE under θ . Thus, under θ there is a deviator from the specified strategy. Since a is θ -maximal for 1 and 2, the deviator must be 3. Take 3's θ -optimal deviation under θ , and a SPE outcome following this deviation under ϕ . Since the deviation cannot be profitable under ϕ , the SPE outcome corresponding 3's deviation must be a or b. Suppose first that this outcome is a SPE under θ as well. Since the deviation must be profitable under θ , it must be that the SPE is b. But, importantly, since b is top ranked by 1 and 2, and b is the best outcome 3 can achieve under θ , it must be that b is a SPE outcome of the whole game under θ as well. This contradicts the assumption that Γ implements f^{B} and $f^{\rm B}(\theta) = c$. Suppose then that the SPE outcome corresponding 3's first deviation under ϕ is not a SPE outcome under θ . Then there is a second deviation by some player under θ from the strategy implementing the SPE outcome (a or b) of the deviation. Again, since a and b are θ -maximal for 1 and 2, the deviator must be 3. As in the case of the first deviation, either b is a SPE outcome of the second deviation under both ϕ and θ , contradicting the assumption that $f^{\rm B}$ is implementable and $f^{B}(\theta) = c$, or there is another deviation from a SPE under ϕ . Since the implementing game is of finite length, there must be a finite stage after which no further deviations take place. But then the consecutive θ -optimal choices made by 3 finally implement b under θ . Therefore, being top ranked by 1 and 2, b must be an SPE outcome of the whole game under θ . Thus, f^{B} cannot be implemented.

To summarize the argument made in Example 1, observe that there is *only one* sequence pair $(h, d)(a, \phi, \theta : D)$ meeting $\alpha(i)$ –(iii): $h(a, \phi, \theta : D) = (3)$ and $d(a, \phi, \theta) = (a, b)$. This sequence pair has length K = 0 and satisfies $d(1) \in M^j(\theta, A)$ for $j \neq h(0)$. Under these circumstances $a \in f(\phi) \setminus f(\theta)$ implies $f(\theta) = b$. This observation is valid more generally. We now construct a condition that builds on this idea.

Before expressing our characterization, we need to develop some notation. For any θ , find set $J(\theta, a, D)$ such that

$$J(\theta, a, D) = \{j : a \notin M^J(\theta, D)\}.$$

Note that if f is unanimous and $a \notin f(\theta)$, then $J(\theta, a, D)$ is non-empty. On the other hand, if $J(\theta, a, D)$ is a singleton set and $a \in f(\phi) \setminus f(\theta)$, then there is a priori determined player, namely j such that $\{j\} = J(\theta, a, D)$, who must deviate under θ from the path implementing a under ϕ . Note that if $J(\theta, a, D)$ is singleton set, then NVP would imply that $a \in f(\theta)$.

Definition 6 (*Condition* α^*). There is a set of outcomes *D*, a collection of $\alpha(i)$ –(iv)-sequences $\{(h^*, d^*)(a, \phi, \theta : D)\}_{a \in f(\phi) \setminus f(\theta)}$, and a collection of outcome sets $\{C_i(a, \phi)\}_{a \in f(\phi), i \in I}$ such that $D \supseteq f(\Theta)$, $C_i(a, \phi) \subseteq L_i(a, \phi)$ for all *i*, and $h^*(k')(a, \phi, \theta : D) = i$ for all k' = 0, ..., k implies $d^*(k)(a, \phi, \theta : D) \in C_i(a, \phi)$ for all $\theta \in \Theta$. Moreover, if $J(\theta, a, D) = \{j\}$, the length of $(h^*, d^*)(a, \phi, \theta : D)$ is zero, and $b \in M^j(\theta, C_i(a, \phi)) \cap [\cap_{i \neq j} M^i(\theta, D)]$, then $b \in f(\theta)$.

In words, there is set D, the domain of the implementing mechanism, and, if $a \in f(\phi)$, set $C_i(a, \phi)$ that defines for any *i* his feasible set if other players stick in their ϕ -equilibrium strategy that implements a. If $a \notin f(\theta)$, then there is a sequence of profitable deviations under θ from the equilibrium path implementing a under ϕ . Such sequence is described by $(h^*, d^*)(a, \phi, \theta : D)$. If all potential deviation sequences are induced by a single player, *j*, and his θ -maximal outcome

in $C_j(a, \phi)$ set is maximal also for all other players in D, then the resulting outcome, b, must be a θ -equilibrium, and consequently $b \in f(\theta)$. Since, all potential deviation sequences are, in fact, induced by a single player if all $\alpha(i)$ –(iv) sequences are of length 0 (this is proven in Lemma 2), we are permitted to use the latter property in the statement of the condition.

Note that Condition α^* implies Condition α : The first sentence of the condition requires the existence of a $\alpha(i)$ -(iv)-sequence whenever $a \in f(\phi) \setminus f(\theta)$. Thus without the second sentence, Condition α^* would be equal to α (as we can always choose $C_i(a, \phi) = L_i(a, \phi)$ for all *i*). Note that α^* is weaker than the combination of α and NVP since α^* restricts choice rules only in a *subset* of those cases where n - 1 players agree on a top ranked alternative. In other words, α^* demands that *f* gives players a limited degree of veto-power. It can easily be seen that unanimity and α^* do not imply NVP and α either.

While α plays the same role of Maskin monotonicity in the context of SPE implementation, α^* now plays the role of $\mu(i)$ –(ii) of Moore–Repullo [15], by restricting a choice rule in a subset of those cases where all but one player agree on a particular top ranked alternative. Analogously, we show that α^* is a necessary condition for SPE implementation.

Lemma 2. Choice rule f is SPE implementable only if it satisfies Condition α^* .

Proof. Let Γ implement f, and choose D = g(S). Then $f(\Theta) \subseteq D$. The proof proceeds as follows: The first part of shows that the lemma holds for certain $\alpha(i)$ -(iii)-sequences $\{(h^*, d^*)(a, \phi, \theta : D)\}_{a \in f(\phi) \setminus f(\theta)}$ such that if $J(\theta, a, D) = \{i\}$, $h^*(k)(a, \phi, \theta : D) = i$ for all k = 0, ..., K, and $b \in M^i(\theta, C_i(a, \phi)) \cap [\cap_{j \neq i} M^j(\theta, D)]$, where $C_i(a, \phi)$ comprises all outcomes i can reach under ϕ by unilaterally deviating from the strategy implementing a, then $b \in f(\theta)$. The second part shows that if one cannot find $(h^*, d^*)(a, \phi, \theta : D)$ meeting also $\alpha(iv)$ with length K > 0, then necessarily $h^*(k)(a, \phi, \theta : D) = i$ for all k = 0, ..., K. By the first part, this implies that if there is no $(h^*, d^*)(a, \phi, \theta : D)$ meeting $\alpha(i)$ -(iv) with length K > 0, $\{i\} = J(\theta, a, D)$, and $b \in M^i(\theta, C_i(a, \phi)) \cap [\cap_{i \neq i} M^j(\theta, D)]$, then $b \in f(\theta)$.

Part 1: Take (a, ϕ) such that $a \in f(\phi)$, and find $s \in \overline{SPE}(\Gamma, \phi)$ such that g(s) = a. Suppose that $a \notin f(\theta) = SPE(\Gamma, \theta)$. Then there is the first node y_0 immediately succeeding the node where some player, h_0 , makes a unilateral single shot deviation from *s* that induces only such θ -equilibria that he strictly θ -prefers over *a*. Choose $d_0 = g(s) = a$ and $d_1 = g(s : y_0)$. If $d_1 \in SPE(\Gamma(y_0), \theta)$, then choose K = 0. If $d_1 \notin SPE(\Gamma(y_0), \theta)$, then repeat the procedure to obtain y_1 , h_1 and d_2 , and so forth.

Since Γ has finitely many stages, this procedure generates sequences (h_0, \ldots, h_K) and (d_0, \ldots, d_{K+1}) that represent the successive unilateral, θ -profitable deviations from *s* by players h_0, \ldots, h_K from the paths implementing $(a =)d_0, d_1, \ldots, d_K$, i.e. there are histories y_0, \ldots, y_K such that $g(s : y_k) = d_k$ and h_k deviates from the path $\bar{y}(s : y_k)$. Then $((h_0, \ldots, h_K), (d_0, \ldots, d_{K+1}))$ constitutes an $\alpha(i)$ -(iii)-sequence $(h, d)(a, \phi, \theta : D)$ that meets $\alpha(i)$ -(iii).

Let $HD(a, \phi, \theta : \Gamma)$ comprise all $\alpha(i)$ -(iii)-sequences $(h, d)(a, \phi, \theta : D)$ derived from Γ in the above described way. Pick sequence $(h^*, d^*)(a, \phi, \theta : D) \in HD(a, \phi, \theta : \Gamma)$ such that $d^*(K + 1)(a, \phi, \theta : D) \in \bigcap_{i \neq j} M^j(\theta, D)$ and $h^*(k)(a, \phi, \theta : D) = i$ for all k = 0, ..., K if and only if all $\alpha(i)$ -(iii)-sequences in $HD(a, \phi, \theta : \Gamma)$ have this property.

Recall that $D(\emptyset, s_{-i})$ comprises all outcomes *i* can reach under ϕ by unilaterally deviating from the strategy implementing *a*, i.e.

$$D(\emptyset, s_{-i}) = \{b : g(s'_i, s_{-i}) = b \text{ for some } s'_i \in S_i\}.$$

Choose

 $C_i(a, \phi) = D(\emptyset, s_{-i}).$

Since unilateral deviation cannot improve *i*'s ϕ -payoff when other players adhere their ϕ -equilibrium strategy s_{-i} , $C_i(a, \phi) \subseteq L_i(a, \phi)$. Thus, by construction $h^*(k')(a, \phi, \theta' : D) = i$ for $k' = 0, \ldots, k$ implies $d^*(k)(a, \phi, \theta' : D) \in C_i(a, \phi)$, for all $\theta' \in \Theta$.

Suppose now that $a \notin f(\theta)$, $J(\theta, a, D) = \{i\}$, $h^*(k)(a, \phi, \theta : D) = i$ for all k = 0, ..., K, and $d^*(K + 1)(a, \phi, \theta : D) \in \bigcap_{j \neq i} M^j(\theta, D)$. Then, since $a \in M^j(\theta, D)$ for all $j \neq i$, it must be profitable for *i* to deviate from *s* under θ . Since $h^*(k)(a, \phi, \theta : D) = i$ for all k = 0, ..., K, all the $\alpha(i)$ -(iii)-sequences in $HD(a, \phi, \theta : \Gamma)$ are unilaterally induced, and hence no sequence of *i*'s deviations is followed by *j*'s deviation, for $j \neq i$, i.e. any further deviation by any *j* must induce an θ -equilibrium that *j* does not prefer. This implies that *i* can induce any outcome in $C_i(a, \phi)$ as long as he himself does not deviate from the path implementing the outcome. Thus the strategy that induces outcome $b \in M^i(\theta, C_i(a, \phi))$ is optimal for him. If, in addition, $b \in \bigcap_{i \neq j} M^j(\theta, D)$, then no *j* wants to deviate at any deviation stage of *i*, either. This implies that $b \in SPE(\Gamma, \theta)$. Thus $b \in f(\theta)$.

Part 2: We still need to show that if all elements of $HD(a, \phi, \theta : \Gamma)$ meeting $\alpha(i)$ -(iv) have length K = 0, then $h^*(k)(a, \phi, \theta : D) = i$ for all k = 0, ..., K. To see this, suppose that $h^*(\bar{k})(a, \phi, \theta : D) = j \neq i$ for some \bar{k} . We show that then there is $(h, d)(a, \phi, \theta : D)$ meeting $\alpha(i)$ -(iv) such that K > 0 (from this on, write (h^*, d^*) instead of $(h^*, d^*)(a, \phi, \theta : D)$). If $d^*(K + 1) \notin \bigcap_{j \neq i} M^j(\theta, D)$, then (h^*, d^*) itself meets $\alpha(i)$ -(iv) with K > 0. Thus suppose that $d^*(K + 1) \in \bigcap_{j \neq i} M^j(\theta, D)$.

Since Γ has finitely many stages, and in this case our selection rule does not restrict the choice of (h^*, d^*) in $HD(a, \phi, \theta : \Gamma)$, there is no loss in assuming that any $\alpha(i)$ –(iii)-sequences (h, d) in $HD(a, \phi, \theta : \Gamma)$ that agrees with (h^*, d^*) until stage \bar{k} , satisfies h(k) = i for any $k > \bar{k}$.

If k = K - 1, then (h^*, d^*) meets $\alpha(iv)$ (note that necessarily $h^*(K) = i$). Thus suppose that $\bar{k} < K - 1$. Find the node y immediately following $h^*(\bar{k})$'s deviation on the path implementing $d^*(\bar{k})$, to the path implementing $d^*(\bar{k}+1)$. Then $d^*(\bar{k}+1) = g(s:y)$ and $s: y \in \overline{SPE}(\Gamma(y), \phi)$. Recall that $D(y, s_{-i})$ comprises all outcomes *i* can reach in subgame $\Gamma(y)$ by unilaterally deviating from *s*. Since a unilateral deviation cannot improve *i*'s ϕ -payoff when other players adhere their ϕ -equilibrium strategy s_{-i} in subgame $\Gamma(y)$, $D(y, s_{-i}) \subseteq L_i(d^*(\bar{k}+1), \phi)$.

Since $h^*(k) = i$ for all $k = \bar{k} + 1, ..., K$, it must be that *i* deviates from *s* after *y*. Since $h^*(k) = i$ for all $k > \bar{k}$ for all $\alpha(i)$ -(iii)-sequences in $HD(a, \phi, \theta : \Gamma)$, no sequence of *i*'s deviations is followed by *j*'s deviation, for any $j \neq i$. Thus *i*'s strategy that induces outcome $b \in M^i(\theta, D_i(y, s_{-i}))$ is optimal for him. Since also $b \in \bigcap_{i \neq j} M^j(\theta, D)$, then no *j* wants to deviate at any deviation stage of *i*, either. This implies that $b \in SPE(\Gamma(y), \theta)$. Since $d(\bar{k} + 1) \notin SPE(\Gamma(y), \theta)$, *i*'s deviation must be profitable, or $b \notin L_i(d(k), \theta)$. Thus $b \in L_i(d(\bar{k} + 1), \phi) \setminus L_i(d(\bar{k} + 1), \theta)$ which implies that sequence $\{(h^*, d^*)(0), \ldots, (h^*, d^*)(\bar{k} + 1), b\}$ meets $\alpha(i)$ -(iv) under $(a, \phi, \theta : D)$. \Box

We have seen that Condition α^* and unanimity are necessary for SPE implementation. We now show that they are also *sufficient*. The next theorem summarizes the results regarding necessity, and provides a constructive proof of sufficiency.

To state the constructive proof for sufficiency, we identify a collection of sequences $\{(h^*, d^*)(a, \phi, \theta : D)\}_{a \in f(\phi) \setminus f(\theta)}$ used in Condition α^* , and employ it to construct a new collection of α -sequences. For any $j \in J(\theta, a, D)$, develop a sequence $(h^j, d^j)(a, \phi, \theta : D)$ by first constructing a cycle of players in $J(\theta, a, D) = \{j[1], \ldots, j[m]\}$, where each player acts as a "dummy mover"

(does not change the outcome), and then adding this cycle on the bottom of $(h^*, d^*)(a, \phi, \theta : D)$. More precisely, for any $k \in \{1, ..., m\}$, choose ¹⁴ (drop the $(a, \phi, \theta : D)$ -term in the description of sequences)

$$\begin{aligned} h^{j(k)}(0) &= j[(k+0) \mod m], & d^{j(k)}(0) = a, \\ \vdots & \vdots \\ h^{j(k)}(m-1) &= j[(k+m-1) \mod m], & d^{j(k)}(m-1) = a, \\ h^{j(k)}(m) &= a = h^*(0), & d^{j(k)}(m) = d^*(0), \\ \vdots & \vdots \\ h^{j(k)}(K+m) &= h^*(K), & d^{j(k)}(K+m) = d^*(K), \\ d^{j(k)}(K+m+1) &= d^*(K+1). \end{aligned}$$

Thus (h^j, d^j) is derived from (h^*, d^*) such that $h^j(\cdot)$ first rotates around elements in $J(\theta, a, D)$, starting from *j*, during which $d^j(\cdot)$ is kept fixed in *a*. The cycle takes $m = #J(\theta, a, D)$ stages. From stage *m* onwards, $(h^j, d^j)(k)$ coincides with $(h^*, d^*)(k - m)$.

Since $a \notin M^j(\theta, D)$ for any $j \in J(\theta, a, D)$, $(h^j, d^j)(a, \phi, \theta : D)$ is an $\alpha(i)$ -(iv)-sequence with length $K^j > 1$ if $(h^*, d^*)(a, \phi, \theta : D)$ is of length K > 0, or if $J(\theta, a, D)$ contains more than one element, or if $d^*(1) \notin \bigcap_{i \neq h^*(0)} M^i(\theta, D)$. If the converse is true, i.e. $(h^*, d^*)(a, \phi, \theta : D)$ is of $K = 0, J(\theta, a, D)$ is singleton, and $d^*(1) \in \bigcap_{i \neq h^*(0)} M^i(\theta, D)$, then the unique $(h^j, d^j)(a, \phi, \theta : D)$ is an $\alpha(i)$ -(iii)-sequence. The following remark collects these observations.

Remark 1. If $a \in f(\phi) \setminus f(\theta)$, then either all $(h^j, d^j)(a, \phi, \theta : D)$, $j \in J(\theta, a, D)$, meet $\alpha(i)$ -(iv) and have length $K^j > 0$, or no-one does. In the latter case, the unique $(h^j, d^j)(a, \phi, \theta : D)$ satisfies $j = h^j(0) = h^j(1)$ and $a = d^j(0) = d^j(1) \in \bigcap_{i \neq j} M^i(\theta, D)$ and $d^j(2) \in \bigcap_{i \neq j} M^i(\theta, D)$.

Theorem 2. Let $n \ge 3$. Then a choice rule f is SPE implementable if and only if it satisfies α^* and unanimity w.r.t. some $D \supseteq f(\Theta)$.

Proof. Necessity of α^* and unanimity follow from Theorem 1, and Lemmata 1 and 2. We now prove that they are also sufficient. We use a version of MR (employed also by AS) stage mechanism, denoted by Γ^{MR} . It is shown that if *f* satisfies α^* and unanimity, then $SPE(\Gamma^{MR}, \theta) = f(\theta)$ for all $\theta \in \Theta$.

Assume that f satisfies α^* and unanimity w.r.t. $D \supseteq f(\Theta)$ and $\{C_i(a, \phi)\}_{a \in f(\phi), i \in I}$, and $\{(h^*, d^*)(a, \phi, \theta)\}_{a \in f(\phi) \setminus f(\theta)}$ (drop the argument D from the description). Suppose that θ is the true state.

Stage 0. Each player $i \in I$ announces a triplet $(a^i, \phi^i, n^i) \in A \times \Theta \times \{0, 1, \ldots\}$.

- (a) If $(a^i, \phi^i) = (a, \phi)$, for all $i \in I$, and $a \in f(\phi)$, then implement a.
- (b) If $(a^i, \phi^i) \neq (a^j, \phi^j) = (a, \phi)$, for all $j \in I \setminus \{i\}$, and $a \in f(\phi)$, then, if $a^i = a$, move to stage 1 with sequence $(h^i, d^i)(a, \phi, \phi^i)$, and if $a \neq a^i \in C_i(a, \phi)$, implement a^i .
- (c) In all other cases, if $n^i > n^j$ for all $j \in I \setminus \{i\}$, then *i* chooses an outcome in *D*. Implement this outcome.¹⁵

¹⁴ Where (x) mod m denotes the congruent class of $x \in \mathbb{N}$ modulo $m \in \mathbb{N}$.

¹⁵ Break the ties in favor of the player with the lowest index.

Stage k = 1, ..., K: Each player $i \in I$ announces a number $n_k^i \in \{-1, 0, 1, ...\}$.

- (a) If $n_k^i = -1$ for at least n 1 players, then $h^i(k 1)(a^i, \phi, \phi^i)$ chooses an outcome in *D*. Implement this outcome.
- (b) If $n_k^i = 0$ for all $i \in I$ then implement $d^i(k)$.
- (c) If $n_k^i = 0$ for all $i \in I \setminus \{h^i(k)\}$, and $n_k^{h(k)} \neq 0$ then move to the stage k + 1, or, in case k = K, implement d(K + 1).
- (d) In all other cases, if $n_k^i > n_k^j$, for all $j \in I \setminus \{i\}$ then *i* chooses an outcome in *D*. Implement this outcome.

First we show that $f(\theta) \subseteq SPE(\Gamma^{MR}, \theta)$. Take any (θ, a) such that $a \in f(\theta)$. Consider the following strategy profile: at stage 0 all players announce (θ, a) , and at stages $k = 1, \ldots, K$ all players announce 0. To see why this is an equilibrium, consider the subgame starting at stage k. Only $h^i(k)$'s deviation can change the resulted outcome (which is $h^i(k-1)$'s top ranked alternative in D). Since θ is true state and $d^i(k)R_{h^i(k)}(\theta)d^i(k+1)$ by definition, a deviation is not profitable. Hence, a unanimous announcement 0 at stage k is an equilibrium, for any $k = 1, \ldots, K$. At stage 0 a unilateral deviation by any other player does not change anything. Consequently, the described strategy is an equilibrium and $a \in SPE(\Gamma^{MR}, \theta)$ if $a \in f(\theta)$.

Now we show that $SPE(\Gamma^{MR}, \theta) \subseteq f(\theta)$. Suppose, to the contrary, that $b \in SPE(\Gamma^{MR}, \theta) \setminus f(\theta)$. First we ask what are the possible SPE strategies at stage 0. Either at least n - 1 players announce (ϕ, a) , or at least three players, i, j, and k disagree on their announcements $(\phi^i, a^i), (\phi^j, a^j)$, and (ϕ^k, a^k) . In the latter case, since the integer game does not knock out b, it must satisfy $b \in \bigcap_{j \in I} M^j(\theta, D)$. By unanimity $b \in f(\theta)$, a contradiction. So at least n - 1 players must announce (ϕ, a) such that $a \in f(\phi)$. There are two cases to consider: Either n players announce (ϕ, a) , and b = a, in which case $a \notin f(\theta)$, or n - 1 players announce (ϕ, a) , and some i announces $(\phi^i, a^i) \neq (\phi, a)$. Note that in the latter case i must belong to $J(\theta, a, D)$ since otherwise, as the integer game does not knock out equilibrium, $b \in \bigcap_{j \neq i} M^j(\theta, D)$, which would imply by unanimity that $b \in f(\theta)$. We seek a contradiction by showing that in both cases either unanimity or Condition α^* imply $b \in f(\theta)$.

Recall that $a \in f(\phi) \setminus f(\theta)$ implying that $(h^*, d^*)(a, \phi, \theta)$ and hence $(h^i, d^i)(a, \phi, \theta)$ does exist). The proof proceeds in two steps. First we show that if $(h^i, d^i)(a, \phi, \theta)$ meets $\alpha(i)$ -(iv), and the integer game at stage 0 does not knock out an equilibrium outcome b, then $b \in f(\theta)$ by unanimity. Then we show that if $(h^i, d^i)(a, \phi, \theta)$ does not meet $\alpha(i)$ -(iv), and the integer game at stage 0 does not knock out equilibrium outcome b, then $b \in f(\theta)$ by Condition α^* .

Case 1: $(h^i, d^i)(a, \phi, \theta)$ meets $\alpha(i)$ -(iv). ¹⁶

By Remark 1, one can associate sequence $(h^i, d^i)(a, \phi, \theta)$ meeting $\alpha(i)$ -(iv), or (h^i, d^i) for short, with length $K^i = K + 1 > 0$ for any player $i \in J(\theta, a, D)$. Consider the consequences of $i \in J(\theta, a, D)$ announcing $(a, \phi^i) = (a, \theta)$ at stage 0, and inducing (h^i, d^i) -subgame. Solve the subgame backwards.

At stage K^i there are two possible equilibrium outcomes: either n-1 players choose -1 which results in outcome in $M^{h^i(K^i-1)}(\theta, D)$, or $h^i(K^i)$ chooses 1 while others choose 0, in which case, since the integer game is not triggered, $d^i(K^i+1) \in \bigcap_{i \neq h^i(K^i)} M^j(\theta, D)$ is implemented.

At stage $K^i - 1$, either n - 1 players choose -1 which results in outcome in $M^{h^i(K^i-2)}(\theta, D)$, or $h^i(K^i)$ chooses 1 while others choose 0, in which case, since the integer game is not triggered, $d^i(K+1) \in \bigcap_{j \neq h^i(K^i-1)} M^j(\theta, D)$ is implemented. Since since $h^i(K^i) \neq h^i(K^i-1)$, the

¹⁶ An analogous proof of this case can be found in Abreu and Sen [2] or in Moore and Repullo [14].

latter case implies $d^i(K+1) \in \bigcap_{j \in I} M^j(\theta, D)$. Thus in all cases an outcome in $M^{h^i(K^i-2)}(\theta, D)$ is implemented.

The same applies inductively to all stages k = 0, ..., K, and hence an outcome in $M^{h^i(k)}(\theta, D)$ is implemented at any stage k. In particular an outcome in $M^{h^i(0)}(\theta, D) = M^i(\theta, D)$ is implemented. Thus any $i \in J(\theta, a, D)$ can guarantee an outcome in $M^i(\theta, D)$ by choosing $(h^i, d^i)(a, \phi, \theta)$ at stage 0.

In equilibrium, therefore, either no $i \in J(\theta, a, D)$ profits from inducing an element in $M^i(\theta, D)$, or i induces $b \in M^i(\theta, D)$. In the former case, $a \in \bigcap_{i \in J(\theta, a, D)} M^i(\theta, D)$. But this contradicts the definition of $J(\theta, a, D)$. In the latter case, since b is not knocked out by the integer game at stage 0, it must be that also $b \in \bigcap_{i \neq i} M^j(\theta, D)$. But by unanimity, $b \in f(\theta)$, a contradiction.

Case 2: $(h^i, d^i)(a, \phi, \theta)$ does not meet $\alpha(i)$ -(iv).

By Remark 1, $i = h^i(0) = h^i(1)$ and $a = d^i(0) = d^i(1) \in \bigcap_{j \neq i} M^j(\theta, D)$ and $d^i(2) \in \bigcap_{j \neq i} M^j(\theta, D)$. If *i* announces $(a, \phi^i) = (a, \theta)$ at stage 0, and induces (h^i, d^i) -subgame, outcome $d^i(1)$, $d^i(2)$, or some unanimously agreed outcome *c* becomes implemented. In the third case, $c \in f(\theta)$. Thus suppose that $d^i(1)$ or $d^i(2)$ becomes implemented. Since $d^i(2)P_i(\theta)d^i(1)$, and $d^i(2) \in C_i(a, \phi)$, *i* chooses (h^i, d^i) over outcome *b* in $C_i(a, \phi)$ only if $d^i(2)$ is maximal in $C_i(a, \phi)$. The integer game does not knock out the equilibrium only if the induced outcome *b* or $d^i(2)$, say *b*, satisfies $b \in M^j(\theta, C_i(a, \phi)) \cap [\bigcap_{j \neq i} M^j(\theta, D)]$. But then, by Condition $\alpha^*, b \in f(\theta)$, a contradiction. \Box

As usual, the foregoing proof on sufficiency is constructive. We employ a version of the stage mechanism introduced by MR (denoted by Γ^{MR}) and show that any choice rule satisfying Condition α^* and unanimity can be SPE implemented by this mechanism. Hence, we show that Γ^{MR} is a *canonical* mechanism: it SPE implements *any* choice rule that may potentially be SPE implemented by *any* mechanism.¹⁷

The main property of the proof is that whenever player *i* wishes to deviate under θ from an otherwise unanimous false announcement, a one that would be announced when implementing *a* under ϕ , he would choose "truthfully" sequence $(h^i, d^i)(a, \phi, \theta)$ whenever it meets $\alpha(i)$ –(iv). The reason for this that in any equilibrium such choice would induce his top ranked outcome. But if choice $(h^i, d^i)(a, \phi, \theta)$ would constitute an equilibrium, the resulting outcome must be unanimously top ranked, since the integer game at stage 0 does not knock out the equilibrium. By unanimity, this would imply that the resulting outcome must be in the choice set under θ . Thus the only question is what happens when truthful $(h^i, d^i)(a, \phi, \theta)$ does not meet $\alpha(i)$ –(iv). We show that in this case, if *i* optimally chooses an outcome in $C_i(a, \phi)$. If the choice constitutes an equilibrium, and is not knocked out by stage 0 integer game, the outcome must belong to the choice set under θ , by Condition α^* .

Consider again Example 1. There it holds that $J(\theta, a, \phi)$ is singleton set. Suppose that $a \in f(\phi) \setminus f(\theta)$. Since $a \in D$, $C_3(a, \phi) \subseteq \{a, b\}$, and $\{3\} = J(\theta, a, D)$, we have by Condition α^* that $f(\theta) = \{x\}$ if $x \in M_3(\theta, C_3(a, \phi)) \cap M_1(\theta, D) \cap M_2(\theta, D)$. Since $a \notin f(\theta)$, we must have $b \in D$ and $C_3(a, \phi) = \{a, b\}$. Thus $f(\theta) = \{b\}$. In particular, the Borda rule f^B chooses $f^B(\phi) = \{a\}$ and $f^B(\theta) = \{c\}$, hence it violates Condition α^* .

To the contrary, *plurality with runoff* chooses $f(\phi) = \{a\}$ and $f(\theta) = \{b\}$, and satisfies Conditions α , α^* and unanimity but not NVP. Hence, even if this rule does not satisfy the AS sufficient condition, it can be implemented.

¹⁷ Of course, Γ^{MR} is subject to well known criticism: it is *unbounded* and, if mixed strategies are allowed, it does not satisfy the *best response* property. See [10,11].

The next example highlights the difference between AS sufficient condition and the complete characterization, and the difference between Nash and SPE implementability.

Example 2. Let $I = \{1, 2, 3\}$, $\Theta = \{\phi, \theta\}$, and alternatives $A = \{a, b, c, d, e\}$ are ranked in *strict* order as follows:

$R_1(\phi) = R_1(\theta)$	$R_2(\phi) = R_2(\theta)$	$R_3(\phi)$	$R_3(\theta)$
а	а	е	b
b	b	С	С
С	С	b	е
d	d	а	а
е	е	d	d

The Borda rule f^{B} chooses $f^{B}(\phi) = \{a\}$ and $f^{B}(\theta) = \{b\}$. Clearly f^{B} is not monotonic nor satisfies NVP. Thus it is not Nash implementable nor SPE implementable according to the AS sufficient condition. However, f^{B} does satisfy Condition α^{*} (construct sequences (a, d, e, b) and (3, 1, 3), and choose D = A and $C_{3}(\phi, a) = \{a, d\}$) and unanimity, and is therefore SPE implementable.

3.1. Remarks

First we give a characterization of Nash implementable choice rules by using α -sequences. Note that in the context of Nash equilibrium only one deviation from the equilibrium path is possible. Given this restriction, $\alpha^* (= \alpha + \alpha + \alpha)$ a restricted veto-power condition) has the following form:

Definition 7. There is a set of outcomes $D \supseteq f(\Theta)$, a collection of pairs $\{(h, d)(a, \phi, \theta)\}_{a \in f(\phi) \setminus f(\theta)}$ in $I \times D$, and a collection of sets $\{C_i(a, \phi)\}_{a \in f(\phi), i \in I}$ in D such that $C_i(a, \phi) \subseteq L_i(a, \phi)$ for all i, and $h(a, \phi, \theta) = i$ implies $d(a, \phi, \theta) \in C_i(a, \phi) \setminus L_i(a, \theta)$ for all $\theta \in \Theta$. Moreover, if $b \in M^j(\theta, C_j(a, \phi)) \cap [\cap_{i \neq j} M^i(\theta, D)]$, then $b \in f(\theta)$.

This is equivalent to Moore and Repullo [15] Condition $\mu(i)$ -(ii). Together with unanimity, it fully characterizes Nash implementable choice rules.

Next we study scenarios where the characterization of SPE implementable rules becomes simple. Many interesting choice rules are *neutral*: names of the alternatives do not matter. In such case, if the domain of preferences is large, the choice rule is onto the set of social alternatives. Thus, whenever there is a unanimously top ranked alternative, then this outcome should be selected by the choice rule. Since any outcome is selected by the choice rule in some state, it follows that Condition α implies unanimity. But then, since Condition α^* implies α , it follows that α^* alone is sufficient to SPE implement this choice rule.

Theorem 3. If $A = f(\Theta)$, then f is SPE implementable if and only if it satisfies Condition α^* .

Proof. Necessity follows from Theorem 2. Suppose that *f* satisfies Condition α^* . Let $c \in \bigcap_{i \in I} M^i$ (θ, A) . Then there is ϕ such that $c \in f(\phi)$. By property $\alpha(iii)$, also $c \in f(\theta)$. Thus unanimity is met and, by Theorem 2, *f* can be implemented. \Box *Linearity* is a natural restriction on players' preferences. Formally, preference profile $R_i(\theta)$ is linear if

 $aR_i(\theta)b$ implies $aP_i(\theta)b$, for all $a \neq b$.

If θ induces linear preferences, then $M^i(\theta, D)$ is single valued for any *i* and $D \subseteq A$.

Lemma 3. If Θ consists of linear preferences, then Condition α^* is equivalent to Condition α .

Proof. Clearly α^* implies α . We show the other way. Take *f* meeting Condition α with respect to set *D*. Then, by Theorem 1, there is a family of $\alpha(i)$ -(iv) sequences $\{(h^*, d^*)(a, \phi, \theta : D)\}_{a \in f(\phi) \setminus f(\theta)}$. Choose $C_i(a, \phi) = L_i(a, \phi)$ for each $a \in f(\phi)$, and $i \in I$. Then the families $\{(h^*, d^*)\}$ and $\{C_i(a, \phi)\}$ of sequences and feasible sets meet the properties imposed by α^* .

Fix (a, ϕ, θ) such that $a \in f(\phi) \setminus f(\theta)$, and suppose that f violates α^* at (a, ϕ, θ) , given our choices of $\{(h^*, d^*)\}$ and $\{C_i(a, \phi)\}$. Then $J(\theta, a, \phi) = \{i\}$, the length of $(h^*, d^*)(a, \phi, \theta : D)$ is zero, and there is $b \in M^i(\theta, C_i(a, \phi)) \cap [\cap_{i \neq j} M^j(\theta, D)]$, but $b \notin f(\theta)$. Recall that $J(\theta, a, \phi) = \{i\}$ is equivalent to $a \in \cap_{i \neq j} M^j(\theta, D)$. Since preferences are linear, $M^j(\theta, D)$ is a singleton set, for all j. Thus, for $j \neq i$,

$$\{a\} = \{b\} = M^{j}(\theta, D) = M^{i}(\theta, C_{i}(a, \phi)).$$
(1)

By construction, K = 0, and hence, by $\alpha(iii)$, $h^*(K) = i$, and by $\alpha(i)$ and $\alpha(ii)$, $d^*(0)R_i$ $(\phi)d^*(1)$ and $d^*(1)P_i(\theta)d^*(0)$, respectively. Since $d^*(0) = a$, we have $d^*(1)P_i(\theta)a$ and $d^*(1) \in C_i(a, \phi)$. But this contradicts (1). \Box

If we now replace NVP by unanimity ¹⁸ in the AS characterization, the gap between necessary and sufficient conditions closes when preferences are linear: f is SPE implementable if and only if it meets Condition α and unanimity.

Finally, we argue that property $\alpha(iv)$ is superfluous under linear preferences—it is always satisfied by some pair of sequences (h, d) meeting $\alpha(i)$ –(iii).

Lemma 4. Let Θ consists of linear preferences. If there exists $(h, d)(a, \phi, \theta : D)$ meeting $\alpha(i)$ -(iii), then there exists a sequence $(h', d')(a, \phi, \theta : D)$ meeting $\alpha(i)$ -(iv).

Proof. Given $(a, \phi, \theta : D)$, take (h, d) meeting $\alpha(i)$ -(iii). If $\{d(K + 1)\} \neq M^i(\theta, D)$ for some $j \neq h(K)$, then we are done. Suppose that $\{d(K + 1)\} = M^j(\theta, D)$ for all $j \neq h(K)$. Then, by $\alpha(iii)$ and linearity of preferences, $d(K) \notin M^i(\theta, D)$ for all i. Take some $j \neq h(K)$, and construct (h', d') as follows:

$$\begin{aligned} h'(0) &= h(0), & d'(0) = d(0), \\ \vdots & \vdots \\ h'(K) &= h(K), & d'(K) = d(K), \\ h'(K+1) &= j, & d'(K+1) = d(K), \\ h'(K+2) &= h(K), & d'(K+2) = d(K), \\ & d'(K+3) = d(K+1). \end{aligned}$$

(h', d') meets $\alpha(i)$ -(iv). \Box

¹⁸ If *f* satisfies NVP, then $a \in \bigcap_{j \neq i} M^j(\theta, D)$ implies $a \in f(\theta)$. Thus if $a \in \bigcap_{j \in I} M^j(\theta, D)$, then, by NVP, also $a \in f(\theta)$. Thus unanimity is implied by NVP.

Thus Conditions α and α^- are equivalent when preferences are linear. Summing Theorem 2, and Lemmas 3 and 4, we get the following simpler characterization of SPE implementation under linear preferences.

Theorem 4. If Θ consists of linear orders, then *f* is SPE implementable if and only if it satisfies Condition α^- and unanimity.

Particularly simple characterization is now obtained when preferences are linear and the *f* is onto *A*. In such case, Theorems 3 and 4 imply that Condition α^- is a necessary and sufficient condition for implementation.

Corollary 1. If Θ consists of linear orders, and $A = f(\Theta)$, then f is SPE implementable if and only if it satisfies Condition α^- .

4. Concluding remarks

We have obtained a *complete* characterization of choice rules that can be implemented in subgame perfect Nash equilibrium (SPE implemented) by using multi-stage mechanisms. Our characterization consists of unanimity and Condition α^* , which is based on Condition α of Abreu and Sen [2]. Thus our result fills the gap between Moore and Repullo [14] and Abreu and Sen [2] necessary and sufficient conditions by replacing α with a stronger condition of α^* , and NVP with a weaker condition of unanimity.

The general characterization we have introduced is rather complex. However, a slight domain restriction simplifies the characterization remarkably: with linear preferences α^* becomes equal to α^- , a weakened version of α .

Allowing randomization simplifies the characterization even further. ¹⁹ Vartiainen [23] shows that under linear preferences and randomized mechanisms Condition α^- reduces to the following: If $a \in f(\phi) \setminus f(\theta)$, then there is no $H \subset I$ such that a is bottom ranked under ϕ in D by all $i \in H$ and top ranked under θ by all $i \in I \setminus H$. By Corollary 1 this condition is the necessary and sufficient condition for SPE implementability of neutral voting rules. This form of the characterization is very operationalizable. E.g. SPE implementability of voting rules can be easily studied with it. For example, *no* scoring rule is SPE implementable whereas the top-cycle correspondence, the uncovered correspondence as well as the plurality with runoff can be SPE implemented by a random mechanism.

The question of two-player implementation, which is much harder, has not been touched by this paper. Vartiainen [22] characterizes partially two-player SPE implementable rules, and shows that any strictly individually rational bargaining solution can be SPE implemented (see also [21]).

Acknowledgements

I am indebted to a referee for his/her exceptionally generous comments and, in particular, for pointing out a serious problem in a previous version. I am also grateful to an associate editor for his/her patience and very useful comments. Finally, I want to thank Tomas Sjöström,

¹⁹ Abreu and Matsushima [1], Abreu and Sen [3], and Glazer and Perry [8] show that with randomized mechanisms one can *virtually* implement essentially all choice rules.

Matthew Jackson, Rafael Repullo, Seppo Honkapohja, and Hannu Salonen for their comments and encouragement along the project.

References

- D. Abreu, H. Matsushima, Virtual implementation in iteratively undominated strategies: complete information, Econometrica 60 (1990) 993–1008.
- [2] D. Abreu, A. Sen, Subgame perfect implementation: a necessary and almost sufficient condition, J. Econ. Theory 50 (1990) 285–299.
- [3] D. Abreu, A. Sen, Virtual implementation in nash equilibrium, Econometrica 59 (1991) 997–1021.
- [4] J. Bergin, A. Sen, Extensive form implementation in incomplete information environments, J. Econ. Theory 80 (1998) 222–256.
- [5] S. Brusco, Perfect Bayesian implementation, Econ. Theory 5 (1995) 419-444.
- [6] S. Brusco, Implementation with extensive form games: one round of signaling is not enough, J. Econ. Theory 87 (1999) 356–378.
- [7] B. Dutta, A. Sen, A necessary and sufficient condition for two-person nash implementation, Rev. Econ. Stud. 58 (1991) 121–128.
- [8] J. Glazer, M. Perry, Virtual implementation in backwards induction, Games Econ. Behav. 15 (1996) 27–32.
- [9] M. Herrero, S. Srivastava, Implementation via backward induction, J. Econ. Theory 56 (1992) 70-88.
- [10] M. Jackson, Implementation in undominated strategies, a look at bounded mechanisms, Rev. Econ. Stud. 59 (1992) 757–775.
- [11] M. Jackson, A crash course in implementation course, Soc. Choice Welfare 18 (2001) 655-708.
- [12] E. Maskin, Nash equilibrium and welfare optimality, Rev. Econ. Stud. 66 (1999) 23-38.
- [13] E. Maskin, T. Sjöström, Implementation theory, in: Arrow, Sen, Suzumura (Eds.), Handbook of Social Choice and Welfare 1, 2002, North-Holland.
- [14] J. Moore, R. Repullo, Subgame perfect implementation, Econometrica 56 (1988) 1191-1220.
- [15] J. Moore, R. Repullo, Nash implementation: a full characterization, Econometrica 58 (1990) 1083–1099.
- [16] H. Moulin, Dominance solvable voting schemes, Econometrica 47 (1979) 1337–1351.
- [17] M. Osborne, A. Rubinstein, A Course in Game Theory, MIT Press, Cambridge, MA, 1994.
- [18] T. Sjöström, On the necessary and sufficient conditions for nash implementation, Soc. Choice Welfare 8 (1991) 333–340.
- [20] S. Srivastava, M. Trick, Sophisticated voting rules: the case of two tournaments, Soc. Choice Welfare 13 (1996) 275–289.
- [21] H. Vartiainen, Studies in mechanism design and game theory, Ph.D. Thesis, Department of Economics Research Reports No. 83, University of Helsinki, 1999.
- [22] H. Vartiainen, Two player subgame perfect implementation, Manuscript, University of Helsinki, 2000.
- [23] H. Vartiainen, Subgame perfect implementation of voting rules via randomized mechanisms, Manuscript, Yrjö Jahnsson Foundation, 2002.