International Journal of Game Theory, Vol. 16, Issue 4, page 263–289

Sub-Core Solutions of the Problem of Strong Implementation

By R. Holzman[1]

Abstract: The problem of strong implementation is to determine which social choice correspondences (SCC) can be obtained as the strong equilibrium correspondence of a game form. We introduce the notion of the nucleus of an effectivity function. Under certain conditions, it yields the smallest implementable SCC having that effectivity function. We contrast it with the core, which yields the largest implementable SCC (as shown by Moulin and Peleg), and argue that the smaller solution should be preferred when available. Another result is that the known necessary conditions for implementability are not sufficient, except in the case of (at most) 3 alternatives.

1 Introduction

This paper considers the following implementation problem, introduced by Maskin (1979). Let N be a finite society and let A be a finite set of alternatives. A *social choice correspondence* (SCC) assigns a non-empty subset of A to every possible profile of preferences of the members of N over A. It is interpreted as an ethics of collective decision. One seeks to decentralize the collective decision, by endowing each individual with a finite space of strategies and specifying an outcome in A for every possible choice of strategies. These rules give rise, for any given profile of preferences, to a game in normal form. The appropriate notion of equilibrium for these games depends on the circumstances; here we shall use strong equilibria[2] – hence the word "strong" in the title. Given a SCC, one would like to design the strategy spaces and the outcome function in such a way that for every profile of preferences, the set of strong

[1] R. Holzman, C.O.R.E., Université Catholique de Louvain.

[2] Recall that a strong equilibrium is a strategy combination that offers to no coalition an opportunity to deviate so as to obtain an outcome that is better for all its members; see Definition 2.3.

0020-7276/87/4/263-289 $2.50 © 1987 Physica-Verlag, Heidelberg

equilibrium outcomes of the resulting game will coincide[3] with the set of outcomes prescribed for that profile by the SCC. A SCC is said to be *implementable* if this is possible.

The problem of strong implementation is to determine which SCCs are implementable. Thus, we are interested to know which ethics of collective decision can be sustained by a decentralized scheme, under circumstances that single out strong equilibria as the reasonable outcomes.

Maskin found that in order that a SCC be implementable, it must have the *strong positive association* (SPA) property[4]. This property requires that if an alternative a belongs to the choice set of a profile, and one considers another profile that maintains all personal preferences for a in the first profile (and possibly adds to them), then a should belong to the choice set of the other profile too. The SPA property is a normative constraint on collective decision that is a necessary condition for implementability[5].

To explore implementability further, Moulin and Peleg (1982) fruitfully used the notion of an *effectivity function.* This is a formal object that describes an allocation of decision power. One such function can be associated in a natural way with every SCC[6]. Then, one may approach the implementation problem by fixing an effectivity function and trying to determine the class of implementable SCCs having this effectivity function.

Moulin and Peleg found that for this class to be non-empty, the effectivity function must satisfy two requirements. First, it must be *maximal*; roughly speaking this means that the decision power allocated to a coalition should be as much as is compatible with that allocated to its complement. Second, it must be *stable*; that is, for every profile of preferences, there should exist at least one undominated alternative, i.e., one that no coalition can effectively improve upon.

Thus, it suffices to consider maximal and stable effectivity functions. Let one such function be fixed. The *core correspondence* is the SCC assigning to each profile the set of undominated alternatives. Moulin and Peleg showed that the core correspon-

---

[3] It may be argued that a weaker requirement would be reasonable. Indeed, Dasgupta et al. (1979) consider the case when, for every profile, the set of equilibrium outcomes is non-empty and contained in the set prescribed by the SCC; when the two sets actually coincide for every profile, they speak of *full* implementation. On the other hand, Moulin and Peleg (1982) consider the case when, for every profile, the set of equilibrium outcomes *contains* the set prescribed by the SCC; they term this *partial* implementation, to distinguish it from implementation that is used in the same sense as in this paper. We think that the fact that different writers have considered each of the two opposite inclusions as minimum requirements is, by itself, an indication of the desirability of equality of the two sets.

[4] Maskin called this property "monotonicity". We adopt however the terminology established later by Moulin and Peleg.

[5] As regards the necessity of another property discussed by Maskin, involving veto power, Maskin's statement was inaccurate; see Moulin and Peleg (1982, Remark 6.8).

[6] We shall recall the definitions in Section 3 below.

dence is implementable, and moreover it is the largest implementable SCC in its class (any implementable SCC having the same fixed effectivity function is contained in it).

Little is known, however, about the rest of such a class, that is, about implementable SCCs that are strictly contained[7] in the core correspondence. It is to these sub-core solutions of the implementation problem that we devote the present research. The reason that the core solution is not quite satisfying is that it is generally too large. While some ties must be allowed for in social choice (a tie being displayed by the choice set consisting of more than one alternative), too many of them undermine the role of the SCC as prescribing collective decision. Hence the interest in smaller SCCs.

We shall propose a new concept – the *nucleus* – comparable to the core but in general smaller than the core. If the effectivity function is nucleus-stable, i.e., it yields a non-empty nucleus for every profile of preferences, then a SCC – the nucleus correspondence – obtains. While the core correspondence is the largest implementable SCC, the nucleus correspondence turns out to be the smallest implementable SCC. Indeed, we prove the following.

Theorem A: Let E be a maximal and nucleus-stable effectivity function. Then the nucleus correspondence of E is implementable. Moreover, it is contained in any implementable SCC with effectivity function E.

Thus, the nucleus gives the best solution to the implementation problem, where "best" is viewed in terms of minimizing the choice sets. This solution is available, however, only for nucleus-stable effectivity functions. We are able to characterize these effectivity functions (Theorem 7.3). This is a rich family of effectivity functions, but strictly smaller than the family of (core) stable effectivity functions. We also identify an important family of effectivity functions for which the nucleus and the core always coincide (Theorem 7.5). In this case, the results for the nucleus and for the core combine to yield uniqueness of the solution.

In general, however, there may exist other solutions, and their characterization remains an open question. We know, by the results of Maskin and Moulin-Peleg, that any implementable SCC must have the SPA property and be contained in the core correspondence of its maximal and stable effectivity function. Are these conditions also sufficient for implementability? In general, the answer is negative, as we show by means of a counterexample with 4 alternatives. When there are no more than 3 alternatives, however, the answer is affirmative, as we prove the following.

---

[7] I.e., the choice set is contained in the core for every profile, it is strictly contained for at least

Theorem B: Let there be at most 3 alternatives and let E be a maximal and stable effectivity function. Then any SCC that has the SPA property and is contained in the core correspondence of E is implementable.

Thus, the implementable SCCs for 3 alternatives (or less) are completely characterized. We remark that Theorems A and B, which may not seem related at first sight, are actually proved jointly (for the most part). Indeed, although the nucleus is not mentioned in Theorem B (and the theorem holds irrespective of nucleus-stability), it plays a role in its proof. The nucleus is also used in the construction of the counterexample mentioned above. These remarks point to the powerful and unifying role played by the nucleus in the resolution of the implementation problem.

The paper is organized as follows. Section 2 introduces formally the problem of strong implementation. In Section 3 we recall the definitions and the facts concerning effectivity functions and cores. In Section 4 we define the nucleus and, recalling the SPA property, we establish an interesting relation (Proposition 4.6) that is used later. Section 5 introduces a game form that is used in Section 6 to prove Theorems A and B. In Section 7 we characterize nucleus-stability and give sufficient conditions for the nucleus and core to coincide. Section 8 is devoted to a counterexample to the statement of Theorem B when there are 4 alternatives.

2 Strong Implementation

Let $N = \{1, \ldots, n\}$ be a finite set of individuals and let A be a finite set of alternatives. It is assumed that $|A|$, the cardinality of A, is at least 2. L will denote the set of all linear orders on A, i.e., we consider preferences that are reflexive, complete, transitive and antisymmetric. L^N will denote the set of all profiles. Thus, an element R^N of L^N is an n-tuple of elements of L, with coordinate R^i standing for the preference order of i. Finally, $\mathcal{P}(A)$ denotes the set of all non-empty subsets of A.

Definition 2.1: A social choice correspondence (SCC) is a function $H: L^N \to \mathcal{P}(A)$ which satisfies[8]: for every $a \in A$ there exists $R^N \in L^N$ such that $H(R^N) = \{a\}$.

Definition 2.2: A game form is an $(n+1)$-tuple $G = (X^1, \ldots, X^n; \pi)$, where X^i, $i = 1, \ldots, n$, is a non-empty finite set (the strategy space of i) and π is a function from the Cartesian product $X^N = \otimes_{i \in N} X^i$ onto A (this is the outcome function).

---

[8] In adding this requirement we follow Moulin and Peleg (1982). It is not essential.

For interpretation, a SCC is a desirable ethics of collective decision, while a game form is a mechanism for society to choose an alternative. To connect the two, we introduce the equilibrium notion that is assumed in our context[9] to predict the behaviour of the individuals, given the game form and their preferences.

The following terminology and notation will be used. A *coalition* is a non-empty subset of N. For a coalition T, $X^T = \otimes_{i \in T} X^i$ denotes the Cartesian product of the strategy spaces of the members of T. When an n-tuple $x^N \in X^N$ is under consideration, x^T will denote the restriction of x^N to T ($x^T \in X^T$).

Definition 2.3: Let $G = (X^1, \ldots, X^n; \pi)$ be a game form and let $R^N \in L^N$. $x^N \in X^N$ is a strong equilibrium point (SEP) of G for R^N if for every coalition S and for every $y^S \in X^S$ there exists $i \in S$ such that $\pi(x^N) R^i \pi(y^S, x^{N-S})$. SEO($G$, R^N) denotes the set of strong equilibrium outcomes, i.e., $a \in$ SEO(G, R^N) if there exists a SEP x^N of G for R^N such that $\pi(x^N) = a$.

Definition 2.4: Let H be a SCC and let G be a game form. We say that H is implemented by G if SEO(G, R^N) = $H(R^N)$ for all $R^N \in L^N$. We say that H is implementable if it is implemented by some game form.

The idea is that if H is implemented by G then a planner who views H as desirable, but is constrained to use a decentralized mechanism, can achieve his goal using G. Hence the interest in implementable SCCs.

3 Effectivity Functions and Cores

We maintain the framework of the previous section. $\mathcal{P}(N)$ denotes the set of all coalitions (non-empty subsets of N). $\mathcal{P}^2(A) = \mathcal{P}(\mathcal{P}(A))$.

Definition 3.1: An effectivity function is a function E: $\mathcal{P}(N) \to \mathcal{P}^2(A)$ which satisfies:

$$E(N) = \mathcal{P}(A), \tag{3.1}$$

[9] The individuals are fully informed and cooperation is unrestricted.

$$[B_i \in E(S_i), \quad i = 1, 2, \quad \text{and}$$

$$S_1 \cap S_2 = \emptyset] \Rightarrow B_1 \cap B_2 \in E(S_1 \cup S_2), \quad \text{(superadditivity)} \tag{3.2}$$

$$[B \in E(S) \text{ and } B \subset B^* \subset A] \Rightarrow B^* \in E(S) \quad (A\text{-monotonicity}). \tag{3.3}$$

Thus, an effectivity function amounts to a relation $B \in E(S)$, where B varies over $\mathcal{P}(A)$ and S varies over $\mathcal{P}(N)$. The interpretation of $B \in E(S)$ is that S can force the outcome to be in B. As (3.1)–(3.3) are natural under this interpretation, they are imposed in the definition. Two other natural properties were not included in the definition, because they are logical consequences of it. First:

$$A \in E(S) \quad \text{for all } S \in \mathcal{P}(N). \tag{3.4}$$

To see this, bear in mind that $E(S) \neq \emptyset$ and apply (3.3). Second:

$$[B \in E(S) \text{ and } S \subset T \subset N] \Rightarrow B \in E(T) \quad (N\text{-monotonicity}). \tag{3.5}$$

To see this, assuming that $S \subsetneq T$, use $B \in E(S)$ and $A \in E(T - S)$ (from (3.4)) to conclude that $B \in E(T)$ from (3.2).

Example 3.2: An important family of effectivity functions is obtained as follows. Let p and q be positive probability measures on A and N respectively. That is, we have positive numbers $\{p(a)\}_{a \in A}$ that add up to 1, and $p(B) = \sum_{a \in B} p(a)$ is the probability of B, and similarly for q on N. Then an effectivity function $E = E(p, q)$ is defined by:

$$B \in E(S) \Leftrightarrow q(S) > 1 - p(B). \tag{3.6}$$

The interpretation is the following: each individual i is endowed with a fraction $q(i)$ of veto power, and for each alternative a, an amount $p(a)$ of veto is required to ban it; (3.6) states that S is effective for B if its total endowment of veto power suffices to veto everything outside B. The strict inequality is required to ensure that $E(p, q)$ is indeed an effectivity function. An effectivity function is called *additive* if it can be obtained from appropriate p and q via (3.6). Additive effectivity functions provide virtually all examples to the theory.

Until now, we have treated effectivity functions as abstract objects. We now make the connection with SCCs[10]. For a coalition T, we let L^T denote the set of all profiles of preferences of the members of T.

Definition 3.3: Let H be a SCC, let $S \in \mathcal{P}(N)$ and $B \in \mathcal{P}(A)$. We say that S is α-effective for B if there exists $R^S \in L^S$ such that for all $Q^{N-S} \in L^{N-S}$, $H(R^S, Q^{N-S}) \subset B$. The effectivity function of H, $E = E(H)$, is defined by: $B \in E(S) \Leftrightarrow S$ is α-effective for B.

One can learn much about H from $E(H)$, and in particular $E(H)$ is useful when one inquires into the implementability of H. This line of research was pursued by Moulin and Peleg (1982), whose results we recall here, without proofs.

Definition 3.4: An effectivity function E is maximal if for all $S \in \mathcal{P}(N)$ and $B \in \mathcal{P}(A)$, $B \notin E(S) \Rightarrow A - B \in E(N - S)$.

Proposition 3.5: If H is an implementable SCC then $E(H)$ is maximal.

To state further results, we require the important concept of the core. For a profile R^N, an alternative a and a coalition S, the following notation will be used:

$$\Pr(S, a, R^N) = \{b \in A \mid b \neq a \text{ and } bR^i a \text{ for all } i \in S\}. \tag{3.7}$$

Definition 3.6: Let E be an effectivity function and let $R^N \in L^N$. The core of E with respect to R^N is the set:

$$C(E, R^N) = \{a \in A \mid \text{for no } S \in \mathcal{P}(N) \Pr(S, a, R^N) \in E(S)\}.$$

If $C(E, R^N) \neq \emptyset$ for all $R^N \in L^N$ we say that E is stable and $C(E, \cdot)$ is called the core correspondence of E.

---

[10] This connection may be done, a priori, in several ways, depending on the notion of effectiveness that one wishes to employ. Our choice follows that of Moulin and Peleg (1982); we refer the reader to that paper for a discussion of the possibilities as well as for the fact that all of them coincide for implementable SCCs. See also Remark 3.9 below for another equivalent way to determine the effectivity function of an implementable SCC.

Definition 3.7: Let H and H^* be SCCs. We say that H is a subcorrespondence of H^* (and also that H is contained in H^*) if $H(R^N) \subset H^*(R^N)$ for all $R^N \in L^N$.

Proposition 3.8: If H is an implementable SCC then $E(H)$ is stable and H is contained in the core correspondence of $E(H)$.

By combining Propositions 3.5 and 3.8, we are led to the following classification of implementable SCCs. For an effectivity function E we denote:

$$\mathcal{I}(E) = \{H \mid H \text{ is an implementable SCC and } E(H) = E\}. \tag{3.8}$$

The class $\mathcal{I}(E)$ is non-empty only if E is maximal and stable. Hence the class of all implementable SCCs is the disjoint union of the classes $\mathcal{I}(E)$, when E varies over all maximal and stable effectivity functions. Fixing one such E, we concentrate on the class $\mathcal{I}(E)$. By Proposition 3.8, it consists of subcorrespondences of the core correspondence of E.

Remark 3.9: Let E be a maximal and stable effectivity function, and let H be a subcorrespondence of $C(E, \cdot)$. Then $E(H) = E$.

It follows that the effectivity function E of an implementable SCC H is uniquely determined within the family of maximal and stable effectivity functions, by the condition: H is a subcorrespondence of $C(E, \cdot)$. Hence we can write alternatively:

$$\mathcal{I}(E) = \{H \mid H \text{ is an implementable subcorrespondence of } C(E, \cdot)\}. \tag{3.9}$$

Is this class non-empty for every maximal and stable effectivity function E?

Theorem 3.10: Let E be a maximal and stable effectivity function. Then the core correspondence of E is implementable.

So $\mathcal{I}(E)$ is non-empty and $C(E, \cdot)$ is its greatest element. Are there any other elements? Only one particular example has been given: Moulin and Peleg (1982, Example 2.38). The present work looks more systematically into this question.

4 Nuclei and Strong Positive Association

Like the core, the nucleus will be defined with respect to an effectivity function. We confine ourselves to maximal effectivity functions. For such E, we introduce the following terminology. Let S be a coalition and let B be a non-empty proper subset of the set of alternatives A. The relation $S \in \mathcal{B}(B) - S$ is *blocking* for $B -$ is defined by:

$$S \in \mathcal{B}(B) \Leftrightarrow A - B \in E(S) \Leftrightarrow B \notin E(N - S). \tag{4.1}$$

The equivalence between the second and third terms follows from superadditivity ((3.2)) and maximality. When B is a singleton $\{b\}$, we shall write $S \in \mathcal{B}(b)$ instead of $S \in \mathcal{B}(\{b\})$.

Definition 4.1: Let E be a maximal effectivity function and let $R^N \in L^N$. The nucleus of E with respect to R^N is the set $\mathrm{Nu}(E, R^N) \subset A$ defined by: $a \in \mathrm{Nu}(E, R^N) \Leftrightarrow$ there exist coalitions S_b, one for each $b \in A - \{a\}$, such that

$$S_b, \quad b \in A - \{a\}, \quad \text{are pairwise disjoint,} \tag{4.2}$$

$$\text{for every} \quad b \in A - \{a\} \quad \text{and every} \quad i \in S_b \; aR^i b, \tag{4.3}$$

$$\text{for every} \quad b \in A - \{a\}, \quad S_b \in \mathcal{B}(b). \tag{4.4}$$

If $\mathrm{Nu}(E, R^N) \neq \emptyset$ for all $R^N \in L^N$ we say that E is nucleus-stable and $\mathrm{Nu}(E, \cdot)$ is called the nucleus correspondence of E.

One may interpret the definition as follows: $a \in \mathrm{Nu}(E, R^N)$ if its choice can be defended against any other proposal, in the sense that the members of S_b would veto b (they would want to by (4.3), they would be able to by (4.4)), in such a way that no individual would exercise his veto more than once ((4.2)).

A comparison with the core is in order. For a profile R^N and distinct alternatives a and b let us denote:

$$V_b^a(R^N) = \{i \in N \mid aR^i b\}. \tag{4.5}$$

It is straightforward to see that the core of E with respect to R^N takes the form:

$$a \in C(E, R^N) \Leftrightarrow \text{for all } B \in \mathcal{P}(A - \{a\}), \quad \bigcup_{b \in B} V_b^a(R^N) \in \mathcal{B}(B). \tag{4.6}$$

For $a \in \mathrm{Nu}(E, R^N)$, we require that it be possible to select, for each $b \in A - \{a\}$, a subset S_b of $V_b^a(R^N)$ such that $S_b \in \mathcal{B}(b)$ and different S_b are disjoint. This can be seen as an explicit way of satisfying the right hand side of (4.6). Indeed, if we have such a selection, then for all $B \in \mathcal{P}(A - \{a\})$ $\bigcup_{b \in B} S_b \in \mathcal{B}(B)$ by superadditivity. Thus we have:

Proposition 4.2: Let E be a maximal effectivity function. Then $\mathrm{Nu}(E, R^N) \subset C(E, R^N)$ for all $R^N \in L^N$.

On the other hand, the right hand side of (4.6) may be satisfied without the existence of a selection as required for the nucleus.

Example 4.3: Consider the effectivity function derived from the "proportional veto principle" (Moulin 1981), for 3 alternatives and 4 individuals. Namely, let $|A| = 3$, $|N| = 4$ and let $E = E(p, q)$ where p and q are the uniform probability measures on A and N respectively (see Example 3.2). In terms of blocking, this means that it takes 2 people to block an alternative and it takes 3 people to block a set of 2 alternatives. Consider also the following profile R^N:

1	2	3	4
a	c	b	c
b	a	a	b
c	b	c	a

We have $C(E, R^N) = \{a, b\}$. For instance, $a \in C(E, R^N)$ because: $V_b^a = \{1, 2\}$ has 2 members (hence is blocking for b), $V_c^a = \{1, 3\}$ has 2 members and $V_b^a \cup V_c^a = \{1, 2, 3\}$ has 3 members (hence is blocking for $\{b, c\}$). The position of b is symmetric to that of a. However, $a \notin \mathrm{Nu}(E, R^N)$ since, to satisfy the condition for the nucleus, the selection would have to take 2 members of V_b^a (hence all of it) and similarly for V_c^a, so 1 would participate in both. In fact, $\mathrm{Nu}(E, R^N) = \emptyset$.

The effectivity function in the above example is maximal and stable (this follows from a general result of Moulin (1981)) but not nucleus-stable, as we have seen. We shall return to this topic in Section 7, where we characterize the family of nucleus-

stable effectivity functions. Our result there indicates that this family is considerably smaller than the family of maximal and stable effectivity functions. This limits the applicability of the concept of the nucleus as compared with the core. However, if E is maximal and nucleus-stable (and, a fortiori, stable), we can contrast the nucleus and core correspondences. They are, respectively, the smallest and largest implementable SCCs having effectivity function E – see Theorem A in the introduction and the last part of Section 3 involving the class $\mathcal{I}(E)$. As we argued in the introduction, being the smallest is an attractive property of the nucleus correspondence. And the family of nucleus-stable effectivity functions is rich enough to justify interest in this case.

We turn now to the SPA property and its relevance to our subject.

Definition 4.4: Let H be a SCC. We say that H has the strong positive association (SPA) property if it satisfies: for every $R^N \in L^N$ and every $a \in A$ such that $a \in H(R^N)$, if $R_1^N \in L^N$ satisfies $[aR^i b \Rightarrow aR_1^i b$ for all $i \in N$, $b \in A]$ then $a \in H(R_1^N)$.

Theorem 4.5 (Maskin): Any implementable SCC has the SPA property.

Together with Propositions 3.5 and 3.8, we have three known *necessary* conditions for implementability: the SPA property, maximality of the effectivity function and being contained in the core correspondence of the effectivity function. Whether or not these conditions together *suffice* for implementability has been an open question. An example in Section 8 implies a negative answer. For the special case of 3 alternatives, Theorem B gives an affirmative answer. We remark that Theorem B has a non-vacuous contribution: by applying it we can deduce the implementability of SCCs that are neither the core nor the nucleus correspondence. Indeed, Moulin and Peleg (1982, Example 2.38) constructed a game form with 3 alternatives and 4 individuals that implements a proper subcorrespondence of $C(E, \cdot)$, where E is the same effectivity function as the one considered in our Example 4.3. Our Theorem B puts their ad hoc example into perspective.

We proceed to show the special role of the nucleus with respect to the SPA property.

Proposition 4.6: Let E be a maximal and stable effectivity function. Let H be a subcorrespondence of $C(E, \cdot)$. If H has the SPA property, then $\mathrm{Nu}(E, R^N) \subset H(R^N)$ for all $R^N \in L^N$.

Proof: Let $a \in \mathrm{Nu}(E, R^N)$. Let $\{S_b\}_{b \in A - \{a\}}$ satisfy (4.2)–(4.4). Define a profile Q^N by modifying R^N as follows: for each $b \in A - \{a\}$ and each $i \in S_b$, b is relegated to the bottom of the preference order of i. Since the S_b are disjoint, this can be carried out. For each $b \in A - \{a\}$ we have (see (3.7)):

$$\Pr(S_b, b, Q^N) = A - \{b\} \in E(S_b),$$

since $S_b \in \mathcal{B}(b)$. Therefore $b \notin C(E, Q^N)$. So we have:

$$H(Q^N) \subset C(E, Q^N) \subset \{a\}.$$

Hence $H(Q^N) = \{a\}$. Now (4.3) permits us to apply the SPA property to conclude that $a \in H(R^N)$.

Of course, if E is nucleus-stable then $\mathrm{Nu}(E, \cdot)$ itself is a SCC that has the SPA property. But notice that in Proposition 4.6 we did not assume nucleus-stability. If E is not nucleus-stable, there will be profiles with $\mathrm{Nu}(E, R^N) = \emptyset$, but for the profiles that do have a non-empty nucleus the proposition is still informative: the nucleus elements must be included in any subcorrespondence of $C(E, \cdot)$ if we insist on the SPA property. Thus, the concept of the nucleus is useful even if by itself it does not define a correspondence.

Taking into account Theorem 4.5 and Propositions 3.5 and 3.8, Proposition 4.6 implies:

Corollary 4.7: If H is an implementable SCC then $\mathrm{Nu}(E(H), R^N) \subset H(R^N)$ for all $R^N \in L^N$.

This takes care of the "moreover" part of Theorem A.

5 The Game Form

Starting from a given SCC H, we shall define here a game form $G_H = (X^1, \ldots, X^n; \pi)$ that will be shown to implement H (under appropriate assumptions). By proving this, in the next section, we shall establish Theorems A and B.

The basic structure of the game form is borrowed from Moulin and Peleg (1982). They used it to implement the core correspondence. Our construction modifies theirs so as to suit a SCC H that is a subcorrespondence of the core correspondence. The idea is to restrict the strategy spaces so as to eliminate the undesirable strong equilibrium outcomes (those that are not in the choice set indicated by H for the profile) while maintaining the desirable ones.

Let H be a SCC and let $E = E(H)$ be its effectivity function. A strategy of player i in the game form G_H will be a pair $x^i = (x_0^i, x_1^i)$, where: x_0^i is a function indicating a

set of alternatives for every coalition that i is a member of (but not every such function is permissible; we shall specify below the requirements), and x_1^i is an alternative (this part of the strategy will play a role only within a tie-breaking device).

We proceed to define the strategy spaces formally. Let $i \in N$. Denote:

$$\Sigma^i = \{T \in \mathcal{P}(N) \mid i \in T\}. \tag{5.1}$$

Consider a function $f : \Sigma^i \to \mathcal{P}(A)$. We want f to satisfy:

$$f(T) \in E(T) \quad \text{for all } T \in \Sigma^i. \tag{5.2}$$

Furthermore, we impose the following restriction. Let $T_1, T_2 \in \Sigma^i$ and let $b_1, b_2 \in A$, $b_1 \neq b_2$. The notation $\varphi(T_1, T_2, b_1, b_2)$ will stand for a statement as follows:

$$\begin{aligned} \varphi(T_1, T_2, b_1, b_2) \Leftrightarrow\ & \text{there exist } R^N \in L^N,\ a \in A - \{b_1, b_2\} \\ & \text{such that } a \notin H(R^N) \text{ and}^{11} \\ & T_j \subset V^a_{b_j}(R^N), \quad j = 1, 2. \end{aligned} \tag{5.3}$$

We require that f satisfy:

$$\begin{aligned} & \text{for all } T_1,\ T_2 \in \Sigma^i,\ b_1,\ b_2 \in A,\ b_1 \neq b_2, \\ & [\varphi(T_1, T_2, b_1, b_2) \Rightarrow \text{there exists } j \in \{1, 2\} \text{ such that } b_j \in f(T_j)]. \end{aligned} \tag{5.4}$$

We interrupt the formal definition to offer some interpretation. When i picks a function $f : \Sigma^i \to \mathcal{P}(A)$, he announces, for every coalition he may find himself in, what action the coalition should take ($f(T) = B$ is interpreted as: T should impose B, i. e., veto the alternatives in $A - B$). The manner in which coalitions are actually formed will be specified below; it is determined by the need for agreement among the members as to the action to be taken. The requirement (5.2) says that a player can propose only feasible actions.

To understand (5.4), suppose that the right hand side is violated, that is, $b_j \notin f(T_j)$, $j = 1, 2$. Thus, i announces: "In T_1 I shall propose to veto b_1 and in T_2 I shall propose

---

[11] Recall the notation (4.5).

to veto b_2". Consider for illustration the profile R^N from Example 4.3. We saw that $a \in C(E, R^N)$. To support a as an equilibrium outcome, 1 would want to veto b with 2 against 3 and 4, and at the same time to veto c with 3 against 2 and 4. We also saw that $a \notin \mathrm{Nu}(E, R^N)$, and this was because of this "double role" that 1 had to play. This suggests that in order to destroy undesirable equilibria, and thus to implement a SCC that is in general between the nucleus and the core correspondence, we should prohibit announcements of the type described. $\varphi(T_1, T_2, b_1, b_2)$ describes the circumstances in which we want to impose the restriction: we do not want a to be an equilibrium outcome for R^N, since $a \notin H(R^N)$; the coalitions T_1, T_2 may serve to defend a against potential deviations leading to b_1, b_2 respectively, since $T_j \subset V^a_{b_j}(R^N); j = 1, 2$. It should be borne in mind, however, that once a restriction is introduced because of a certain profile R^N, it is also a constraint in the construction of equilibria for profiles other than R^N. This procedure may conceivably result in destroying desirable equilibria as well, and this is one of the things that we shall have to worry about in the proof.

Back to the formal definition, we let:

$$X_0^i = \{x_0^i \mid x_0^i : \Sigma^i \to \mathcal{P}(A) \text{ satisfies (5.2) and (5.4)}\},$$

$$X^i = X_0^i \otimes A. \tag{5.5}$$

We remark that $X_0^i \neq \emptyset$; indeed, the constant function $x_0^i(T) = A$ is in X_0^i. We have therefore well-defined strategy spaces X^i for $i \in N$.

To define the outcome function π, let $x^N \in X^N$ where $x^i = (x_0^i, x_1^i)$ is the strategy of i, and let $x_0^N = (x_0^1, \ldots, x_0^n)$. For any coalition T, we denote by $Q(T, x_0^N)$ the partition of T into equivalence classes with respect to the equivalence relation: $i \sim j \Leftrightarrow x_0^i(T) = x_0^j(T)$. Consider the sequence of partitions of N defined inductively by:

$$P_0(x_0^N) = \{N\},$$

$$P_{k+1}(x_0^N) = \bigcup_{T \in P_k(x_0^N)} Q(T, x_0^N). \tag{5.6}$$

In words, we split the players into parts according to what they propose for N, then we split the players in each part according to what they propose for the part they are in, and so on. Each partition is a refinement of the preceding one. Hence there exists a first h such that $P_h(x_0^N) = P_{h+1}(x_0^N)$. We call $P(x_0^N) = P_h(x_0^N)$ the *partition associated with* x_0^N. By construction we have:

$$i, j \in T \in P(x_0^N) \Rightarrow x_0^i(T) = x_0^j(T). \tag{5.7}$$

Hence we may define for $T \in P(x_0^N)$:

$$x_0^N(T) = x_0^i(T) \quad \text{for any } i \in T. \tag{5.8}$$

We let:

$$B(x_0^N) = \bigcap_{T \in P(x_0^N)} x_0^N(T). \tag{5.9}$$

To sum up, we obtain the set of alternatives $B(x_0^N)$ by first partitioning N until each part displays unanimity as to the set of alternatives that this part should veto, and then enforcing these vetoes. By superadditivity of E, $B(x_0^N) \neq \emptyset$.

We want the outcome $\pi(x^N)$ to be in $B(x_0^N)$. To take care of the case where $B(x_0^N)$ is not a singleton, we specify a tie-breaking device. First, we endow the set of alternatives A with a group structure, whose operation we denote by $+$. By commutativity and associativity, sums of the form $\sum_{i \in N} a^i = a^1 + \ldots + a^n$ are well-defined. Next, we choose for each $B \in P(A)$ a function ρ_B from A *onto* B. Both the group structure and the functions ρ_B are chosen arbitrarily and are fixed in advance (independent of x^N).

We complete the definition of the outcome function by letting:

$$\pi(x^N) = \rho_{B(x_0^N)} \left(\sum_{i \in N} x_1^i \right). \tag{5.10}$$

Thus, once the set $B(x_0^N)$ is determined, the second coordinates of the players' strategies (which are just elements x_1^i of A) are used to pick one element of $B(x_0^N)$, which is taken as $\pi(x^N)$. In designing this tie-breaking device, we have in mind the following feature. Suppose that x_0^N is given, and let T be a coalition. Then, given x_1^j for $j \notin T$, the elements x_1^i for $i \in T$ may be chosen so as to obtain as $\pi(x^N)$ *any* element of $B(x_0^N)$. One corollary of this is that π is onto A. Indeed, let x_0^N be composed of the constant functions $x_0^i(T) = A$. Then $B(x_0^N) = A$, and x_1^i may be chosen so as to obtain as $\pi(x^N)$ any element of A.

We have thus defined a game form $G_H = (X^1, \ldots, X^n; \pi)$.

6 Proof of Theorems A and B

Since the "moreover" part of Theorem A has already been proved (Corollary 4.7), it remains to prove the first statement of Theorem A and Theorem B. In both cases, we have to prove that a SCC H is implementable. In the case of Theorem A, H is as-

sumed to be the nucleus correspondence of a maximal and nucleus-stable effectivity function. In the case of Theorem B, H is only assumed to have the SPA property and to be contained in the core correspondence of a maximal and stable effectivity function, but it is assumed that there are at most 3 alternatives.

We shall prove that, in each case, the game form $G_H = (X^1, \ldots, X^n; \pi)$ defined in Section 5 implements H. We observe that the effectivity function E that is used in the definition of G_H, taken there as $E = E(H)$, is the same as the one that appears in the assumptions of the theorems; this follows from Remark 3.9 and Proposition 4.2.

To prove that G_H implements H, we show first that $\mathrm{SEO}(G_H, R^N) \subset H(R^N)$ for all $R^N \in \mathrm{L}^N$, and next that $H(R^N) \subset \mathrm{SEO}(G_H, R^N)$ for all $R^N \in L^N$. Since the proof of the first inclusion is the same in both cases (Theorem A and Theorem B), it will be given once for both.

Let $R^N \in L^N$ and let $a \in \mathrm{SEO}(G_H, R^N)$. Let, thus, $x^N \in X^N$ be a strong equilibrium point of G_H for R^N such that $\pi(x^N) = a$. The idea now is to show that, in order to protect a from profitable deviations, there must be in x^N agreement to veto any other alternative that is a possible menace; this agreement must be among players who prefer a to the menace. This is the content of the following claim.

Claim 6.1: Suppose $b \in A - \{a\}$ and $V_b^a(R^N) \neq N$. Then there exists a coalition $S_b \subset V_b^a(R^N)$ such that for all $i \in S_b$, $b \notin x_0^i(S_b)$.

Proof: We recall that a strategy $x^i \in X^i$ is a pair (x_0^i, x_1^i), where $x_0^i \in X_0^i$ and $x_1^i \in A$ (see (5.5)). Let now $V = V_b^a(R^N)$, and consider the combination $(x_0^V, y_0^{N-V}) \in \otimes_{i \in N} X_0^i$, where the components of the members of V are as in the equilibrium and for $i \in N - V$ y_0^i is the constant function $y_0^i(T) = A$.

It must be the case that $b \notin B(x_0^V, y_0^{N-V})$ (see (5.9)). For otherwise there would be a way to choose y_1^i for $i \in N - V$ so that $\pi(x^V, y^{N-V}) = b$ (here x^V is the part of the equilibrium corresponding to the members of V and $y^i = (y_0^i, y_1^i)$ for $i \in N - V$; we are using the feature of the tie-breaking device discussed after formula (5.10) and the fact that $N - V \neq \emptyset$). This would constitute a profitable deviation of the members of $N - V$ from the equilibrium.

Since $b \notin B(x_0^V, y_0^{N-V})$, there is a part T in the partition associated with (x_0^V, y_0^{N-V}) such that $b \notin (x_0^V, y_0^{N-V})(T)$. As $y_0^i(T) = A$ for $i \in N - V$, we must have $T \subset V$. Taking such T as S_b, we satisfy the requirement of the claim.

The above can be done separately for each b that satisfies the assumption of Claim 6.1. Our next observation is that this yields pairwise disjoint S_b's, if we assume that $a \notin H(R^N)$; this will follow from the restriction imposed on strategies involving "double roles".

Claim 6.2: Assume $a \notin H(R^N)$, and suppose $b_1, b_2 \in A - \{a\}$, $b_1 \neq b_2$ and $V_{b_j}^a(R^N) \neq N$, $j = 1, 2$. Let S_{b_j}, $j = 1, 2$, be coalitions as shown to exist in Claim 6.1. Then $S_{b_1} \cap S_{b_2} = \emptyset$.

Proof: Suppose $i \in S_{b_1} \cap S_{b_2}$. Then $b_j \notin x_0^i(S_{b_j})$ for $j = 1, 2$. But $\varphi(S_{b_1}, S_{b_2}, b_1, b_2)$ is true (since we assume $a \notin H(R^N)$ – see (5.3)), hence x_0^i violates (5.4), so $x_0^i \notin X_0^i$. This contradiction proves that $S_{b_1} \cap S_{b_2} = \emptyset$.

We recall now that we started with $a \in \mathrm{SEO}(G_H, R^N)$, and we have to show that $a \in H(R^N)$. Hence we may assume that $a \notin H(R^N)$, and by our claims we obtain pairwise disjoint coalitions S_b, one for each $b \in A - \{a\}$ such that $V_b^a(R^N) \neq N$, with the properties that $S_b \subset V_b^a(R^N)$ and for all $i \in S_b$, $b \notin x_0^i(S_b)$. The last property implies that $S_b \in \mathcal{B}(b)$, since x_0^i has to satisfy (5.2). If we knew that $V_b^a(R^N) \neq N$ for *all* $b \in A - \{a\}$, then the S_b's would testify to the fact that $a \in \mathrm{Nu}(E, R^N)$. From this we could conclude that $a \in H(R^N)$, invoking, in the case of Theorem B, Proposition 4.6. The following claim shows that the existence of b with $V_b^a(R^N) = N$ (for which S_b is not available) does not essentially affect the argument.

Claim 6.3: Suppose that for each $b \in A - \{a\}$, except those b such that $V_b^a(R^N) = N$, a coalition S_b is given, and the S_b are pairwise disjoint and satisfy $S_b \subset V_b^a(R^N)$ and $S_b \in \mathcal{B}(b)$. Then $a \in H(R^N)$.

Proof: We know that H is a subcorrespondence of $C(E, \cdot)$ and has the SPA property (these are explicitly assumed in the case of Theorem B, and are also true when H is the nucleus correspondence as in Theorem A). We proceed to show that $a \in H(R^N)$ just as in the proof of Proposition 4.6. The fact that we do not have S_b for those $b \in A - \{a\}$ that satisfy $V_b^a(R^N) = N$ does not bother us, since such b is Pareto dominated by a in the profile Q^N constructed in the proof, and therefore is not in the core.

This completes the proof of the inclusion $\mathrm{SEO}(G_H, R^N) \subset H(R^N)$.

We have to prove now that $H(R^N) \subset \mathrm{SEO}(G_H, R^N)$ for all $R^N \in L^N$. Assuming $a \in H(R^N)$, we shall construct $x_0^N \in X_0^N = \otimes_{i \in N} X_0^i$ such that:

$$\text{for all } y_0^N \in X_0^N \text{ and all } b \in A - \{a\},$$

$$[y_0^i = x_0^i \text{ for all } i \in V_b^a(R^N)] \Rightarrow b \notin B(y_0^N). \tag{6.1}$$

Clearly, this will suffice, since then for any arbitrary specification of x_1^i for $i \in N$, the combination $x^N \in X^N$, where $x^i = (x_0^i, x_1^i)$, is a SEP of G_H for R^N such that $\pi(x^N) = a$.

The construction of x_0^N will differ for the two theorems. In the case of Theorem A, where H is the nucleus correspondence, we exploit the disjoint structure in the definition of the nucleus. This is done as follows.

Claim 6.4: Assume $a \in \mathrm{Nu}(E, R^N)$. Let S_b, for $b \in A - \{a\}$, satisfy (4.2)–(4.4). Define for $i \in S_b$, $T \in \Sigma^i$:

$$x_0^i(T) = A - \{b\} \quad \text{if } T \supset S_b,$$

$$= A \quad \text{otherwise.}$$

(If $i \notin \bigcup_{b \in A - \{a\}} S_b$, $x_0^i \in X_0^i$ is arbitrary.) Then $x_0^N \in X_0^N$ and (6.1) holds.

Proof: Let $b \in A - \{a\}$. The fact that $x_0^i \in X_0^i$ for $i \in S_b$ follows from $S_b \in \mathcal{B}(b)$ and the observation that (5.4) cannot be violated since only b is ever excluded from $x_0^i(T)$ and either b_1 or b_2 must differ from b.

To prove (6.1), let now $y_0^N \in X_0^N$ coincide with x_0^N for $i \in V_b^a(R^N)$. By (4.3), $y_0^i = x_0^i$ for all $i \in S_b$. Hence, considering the definition of x_0^i for $i \in S_b$, the members of S_b will never be distinguished along the sequence of partitions corresponding to y_0^N, so there will be a part $T \in P(y_0^N)$ with $T \supset S_b$ and therefore $y_0^N(T) = A - \{b\}$. This shows that $b \notin B(y_0^N)$.

In the conditions of Theorem B, we can no longer use the above disjoint scheme. Here we shall construct the equilibrium in the same way as Moulin and Peleg (1982), but we shall need the fact that there are at most 3 alternatives to show that the requirements (5.4) are met. We proceed formally.

Claim 6.5: Under the assumptions of Theorem B, let $a \in H(R^N)$. Define for $i \in N$, $T \in \Sigma^i$:

$$x_0^i(T) = A - \mathrm{Pr}(N - T, a, R^N).$$

(See (3.7); for $T = N$ this reduces to $x_0^i(N) = \{a\}$.) Then $x_0^N \in X_0^N$ and (6.1) holds.

Proof: Since $H(R^N) \subset C(E, R^N)$, we have $a \in C(E, R^N)$, whence $\mathrm{Pr}(N - T, a, R^N) \notin E(N - T)$ for all $T \subsetneq N$. By maximality, $A - \mathrm{Pr}(N - T, a, R^N) \in E(T)$ for all coalitions T (including $T = N$, by (3.1)). This shows that the x_0^i satisfy (5.2).

To establish (5.4), let $T_1, T_2 \in \Sigma^i$ and let $b_1, b_2 \in A$, $b_1 \neq b_2$ and assume $\varphi(T_1, T_2, b_1, b_2)$ holds. Since $a \in x_0^i(T)$ for all $T \in \Sigma^i$, the conclusion of (5.4) holds true if

b_1 or b_2 coincide with a. We may therefore assume that $\{b_1, b_2\} = A - \{a\}$ (as A is assumed to have at most 3 elements). Now $\varphi(T_1, T_2, b_1, b_2)$ means that there exists $Q^N \in L^N$ such that $a \notin H(Q^N)$ and $T_j \subset V_{b_j}^a(Q^N), j = 1, 2$. Since H has the SPA property, the facts that $a \in H(R^N)$ and $a \notin H(Q^N)$ imply the existence of $i_* \in N$ and $j \in \{1, 2\}$ such that $aR^{i}{}_* b_j$ but $b_j Q^{i}{}_* a$. Given that $T_j \subset V_{b_j}^a(Q^N)$, we must have $i_* \in N - T_j$, so by the definition of $x_0^i(T_j)$ we have $b_j \in x_0^i(T_j)$ as required.

To prove (6.1), let $b \in A - \{a\}$ and assume that $y_0^N \in X_0^N$ coincides with x_0^N for $i \in V_b^a(R^N)$. Then for every coalition $T \supset V_b^a(R^N)$ and every i, $i_* \in V_b^a(R^N)$, $b \notin y_0^i(T) = y_0^{i}{}_*(T)$. This permits, as in the proof of Claim 6.4, to argue along the sequence of partitions corresponding to y_0^N, concluding that $b \notin B(y_0^N)$.

This has established the inclusion $H(R^N) \subset \mathrm{SEO}(G_H, R^N)$ for both theorems. Together with the converse that was previously established, we have shown that G_H implements H. Thus, Theorems A and B have been proved.

7 A Characterization of Nucleus-Stability

The family of effectivity functions that arise in the problem of strong implementation consists of the maximal and stable effectivity functions. Theorem A provides an "optimal" solution to the problem for a subfamily of this family, consisting of the maximal and nucleus-stable effectivity functions. Hence it is interesting to know how large this subfamily is, and to be able to determine whether an effectivity function belongs to it.

This is the purpose of the characterization. While nucleus-stability is defined by reference to all possible profiles of preferences, we shall characterize it by means of a property that can be verified by inspection of the effectivity function alone. The characterization is relative to the family of maximal and stable effectivity functions, so it tells us what it takes for a maximal and stable effectivity function to be nucleus-stable as well. We remark that the larger family (maximal and stable) has already been characterized by Peleg (1982), so our relative characterization may be transformed into an absolute one.

Since the definition of the nucleus refers explicitly to the blocking relation determined by the effectivity function (see (4.1)), so will our characterization. We shall use the notation $S \in \mathcal{B}_m(b) - S$ is *minimal blocking* for b – where S is a coalition and b is an alternative, to mean that $S \in \mathcal{B}(b)$ and for all $T \subsetneq S$, $T \notin \mathcal{B}(b)$. The following property will be shown to characterize nucleus-stability.

Definition 7.1: Let E be a maximal effectivity function. We say that E obeys the subtraction rule for blocking (SRB) if it satisfies:

for every two coalitions $S_0 \subset S$,
for every $B \in \mathcal{P}(A)$ such that $1 < |B| < |A|$, and every $b \in B$

$$[S \in \mathcal{B}(B) \text{ and } S_0 \in \mathcal{B}_m(b)] \Rightarrow S - S_0 \in \mathcal{B}(B - \{b\}) \text{ (in particular } S - S_0 \neq \emptyset).$$

The justification of the SRB is that if one takes out of S a coalition that is just "big" enough to block one alternative then the remaining members should be able to carry out the rest of the job (that is, to block the remaining alternatives). One could think of this in terms of dividing the labour between the members of the coalition S. We shall analyse now a simple family of examples, to get an idea as to how often the SRB is satisfied.

Example 7.2: As in Example 4.3, we consider the "proportional veto principle" for 3 alternatives, but here we let the number of individuals vary. Namely, let $|A| = 3$, $|N| = n$ and let $E_n = E(p, q_n)$ where p and q_n are the uniform probability measures on A and N respectively.

By a result of Moulin (1981) E_n is stable for every n. For E_n to be maximal, we need that $n \not\equiv 0 \pmod 3$.

We restrict ourselves to these values of n and inquire into the satisfaction of the SRB by E_n. Assume that $|B| = 2$ (this is the only case to check) and $b \in B$. Then:

$$S \in \mathcal{B}(B) \Leftrightarrow |S| > \frac{2n}{3}, \tag{7.1}$$

$$S_0 \in \mathcal{B}_m(b) \Leftrightarrow |S_0| - 1 < \frac{n}{3} < |S_0|. \tag{7.2}$$

Now, if $n \equiv -1 \pmod 3$ then from the right hand sides of (7.1) and (7.2) one can conclude that $|S - S_0| > n/3$, hence E_n obeys the SRB. On the other hand, if $n \equiv 1 \pmod 3$ then taking $|S| = (2n + 1)/3$ and $S_0 \subset S$ with $|S_0| = (n + 2)/3$ yields a violation of the SRB.

Thus, within the family considered in the example, just "half" obey the SRB. While this is only an example, it does indicate that the SRB is a demanding requirement, yet it is satisfied by many effectivity functions.

We turn now to the characterization theorem.

Theorem 7.3: Let E be a maximal and stable effectivity function. Then E is nucleus-stable if and only if E obeys the SRB.

Proof: We prove first the "only if" part. Assume that the SRB is violated for the coalitions $S_0 \subset S$, the set $B \in \mathcal{P}(A)$ with $1 < |B| < |A|$ and $b \in B$. Denote:

$$S_1 = S - S_0, \qquad S_2 = N - S,$$

$$B_1 = B - \{b\}, \qquad B_2 = A - B.$$

This gives us a partition S_0, S_1, S_2 of N (it is possible that S_1 or S_2 are empty) and a corresponding partition $\{b\}, B_1, B_2$ of A (with all parts non-empty). Since we started with a violation of the SRB, we have:

$$S_0 \in \mathcal{B}_m(b), \tag{7.3}$$

$$S_j \notin \mathcal{B}(B_j), \quad j = 1, 2. \tag{7.4}$$

Consider a profile $R^N \in L^N$ in which the individuals rank the alternatives in accordance with the following scheme[12]:

S_0	S_1	S_2
	b	b
$B_1 \cup B_2$		
	B_2	B_1
b	B_1	B_2

We shall prove that $\mathrm{Nu}(E, R^N) = \emptyset$, contrary to nucleus-stability. Indeed, since $b \notin C(E, R^N)$ by (7.3), b cannot be in the nucleus. Assume therefore that $a \neq b$ and $a \in \mathrm{Nu}(E, R^N)$. By definition of the nucleus, we may find coalitions S_c, one for each $c \in A - \{a\}$, satisfying (4.2)–(4.4). One of these coalitions is S_b. By (4.3) we must have $S_b \subset S_0$, which implies by (4.4) and (7.3) that $S_b = S_0$. By (4.2) it follows that $S_c \subset S_1 \cup S_2$ for every $c \in A - \{a, b\}$.

Assume now that $a \in B_1$ (an analogous argument develops if $a \in B_2$). Then for $c \in B_2$ S_c cannot intersect S_1 in view of (4.3). This implies that $\bigcup_{c \in B_2} S_c \subset S_2$. By (4.2), (4.4) and superadditivity $\bigcup_{c \in B_2} S_c \in \mathcal{B}(B_2)$, which leads to $S_2 \in \mathcal{B}(B_2)$, contradicting (7.4).

---

[12] The rankings within the sets depicted are immaterial.

We prove now the "if" part by induction on $|A|$. For $|A| = 2$ the nucleus coincides with the core, so our assumption that E is stable guarantees that E is nucleus-stable. Assume now that $|A| > 2$ and that $E : \mathcal{P}(N) \to \mathcal{P}^2(A)$ is a maximal and stable effectivity function obeying the SRB.

We remark first that the following fact is implied by the stability of E: for any collection of (possibly empty) subsets S_b of N, one for each $b \in A$,

$$\bigcup_{b \in A} S_b = N \Rightarrow \text{there exists } b \in A \text{ with } S_b \in \mathcal{B}(b). \tag{7.5}$$

This is because if (7.5) fails we have $\bigcap_{b \in A} (N - S_b) = \emptyset$ and $\{b\} \in E(N - S_b)$ for each $b \in A$, and these facts permit the construction of a profile with an empty core (this is a well-known construction in the spirit of the "paradox of voting").

Let now $R^N \in L^N$. We have to show that $\mathrm{Nu}(E, R^N) \neq \emptyset$. Consider, for each $b \in A$, the set of individuals that rank b last:

$$S_b = \{i \in N \mid aR^i b \text{ for all } a \in A\}. \tag{7.6}$$

Applying (7.5), we choose $b_0 \in A$ such that $S_{b_0} \in \mathcal{B}(b_0)$. Next we choose $S_0 \subset S_{b_0}$ such that $S_0 \in \mathcal{B}_m(b_0)$.

We consider now the following restricted sets of individuals and alternatives:

$$\tilde{N} = N - S_0, \quad \tilde{A} = A - \{b_0\}.$$

We remark that $\tilde{N} \neq \emptyset$ by the SRB (consider $S_0 \subset N$ and any B with $1 < |B| < |A|$, $b_0 \in B$). We define now for every $T \in \mathcal{P}(\tilde{N})$:

$$\tilde{E}(T) = \{B - \{b_0\} \mid B \in E(T)\}. \tag{7.7}$$

It is straightforward to verify, using the similar properties for E and the SRB, that (7.7) defines a maximal effectivity function $\tilde{E}$: $\mathcal{P}(\tilde{N}) \to \mathcal{P}^2(\tilde{A})$. Next one observes that the blocking relations $\mathcal{B}$ and $\tilde{\mathcal{B}}$ derived from E and $\tilde{E}$ respectively coincide over the common domains:

$$\text{for all } T \in \mathcal{P}(\tilde{N}),\ B \in \mathcal{P}(\tilde{A})[T \in \tilde{\mathcal{B}}(B) \Leftrightarrow T \in \mathcal{B}(B)]. \tag{7.8}$$

It follows that $\tilde{E}$ obeys the SRB as well. We claim further that $\tilde{E}$ is stable. Indeed, given a profile $\tilde{Q}^{\tilde{N}}$, consider a profile Q^N such that:

(i) for $i \in \tilde{N}$, Q^i is obtained from $\tilde{Q}^i$ by adding b_0 on top of the ranking,

(ii) for $i \in S_0$, Q^i ranks b_0 last.

By the stability of E there exists $a \in C(E, Q^N)$. From (ii) we know that $a \neq b_0$, and from (i) it follows that $a \in C(\tilde{E}, \tilde{Q}^{\tilde{N}})$.

By the induction hypothesis, we know that $\tilde{E}$ is nucleus-stable. Consider now the restriction of R^N to $\tilde{N}$ and $\tilde{A}$ – denote it $\tilde{R}^{\tilde{N}}$. We can find $a \in \mathrm{Nu}(\tilde{E}, \tilde{R}^{\tilde{N}})$. Taking a collection of coalitions that witness this, one for each $b \in \tilde{A} - \{a\}$, together with the coalition S_0 for b_0, we see that $a \in \mathrm{Nu}(E, R^N)$.

This completes the proof of the "if" part of the theorem. We remark that the proof actually indicates a procedure to obtain a nucleus element, by repeated elimination of alternatives and sets of individuals.

It is interesting to note that if we confine ourselves to effectivity functions that do not discriminate among individuals, then the above argument can be replaced by a shorter one that yields a stronger result: *any* core element is in the nucleus. We proceed to formulate and prove this fact.

Definition 7.4: An effectivity function E is anonymous if for every $S, T \in \mathcal{P}(N)$ and every $B \in \mathcal{P}(A)$:

$$[B \in E(S) \text{ and } |T| = |S|] \Rightarrow B \in E(T).$$

Theorem 7.5: Let E be a maximal and anonymous effectivity function that obeys the SRB. Then $C(E, R^N) = \mathrm{Nu}(E, R^N)$ for all $R^N \in L^N$.

Proof: Anonymity implies that to each $B \in \mathcal{P}(A) - \{A\}$ there corresponds an integer $\beta(B)$, $1 \leqslant \beta(B) \leqslant |N|$, such that for all $S \in \mathcal{P}(N)$:

$$S \in \mathcal{B}(B) \Leftrightarrow |S| \geqslant \beta(B). \tag{7.9}$$

The SRB implies now that β is an additive function, that is (denoting $\beta(b) = \beta(\{b\})$), for all $B \in \mathcal{P}(A) - \{A\}$:

$$\beta(B) = \sum_{b \in B} \beta(b). \tag{7.10}$$

(Superadditivity of E yields $\beta(B) \leqslant \sum_{b \in B} \beta(b)$; the inverse inequality is proved by induction on $|B|$, using the SRB.)

Suppose now that $a \in C(E, R^N)$. By (4.6), (7.9) and (7.10), this means that:

$$\text{for all} \quad B \in \mathcal{P}(A - \{a\}) \left| \bigcup_{b \in B} V_b^a(R^N) \right| \geqslant \sum_{b \in B} \beta(b). \tag{7.11}$$

An argument based on a generalization of Hall's "marriage" theorem[13] implies that (7.11) guarantees the existence of pairwise disjoint coalitions S_b, one for each $b \in A - \{a\}$, such that $S_b \subset V_b^a(R^N)$ and $|S_b| = \beta(b)$ for each $b \in A - \{a\}$. This in turn means that $a \in \mathrm{Nu}(E, R^N)$.

This proves that $C(E, R^N) \subset \mathrm{Nu}(E, R^N)$ for all $R^N \in L^N$. Proposition 4.2 supplies the converse, so the proof is complete.

Corollary 7.6. Let E be a maximal, stable and anonymous effectivity function that obeys the SRB. Then there is a unique implementable SCC with effectivity function E: it is the core (and nucleus) correspondence of E.

We see that the concept of the nucleus does not provide a new solution to the implementation problem in the anonymous case. If the SRB is violated we do not have a nucleus correspondence, and if the SRB is satisfied the nucleus correspondence coincides with the core correspondence. Yet the nucleus is useful in other ways: in the former case as a lower bound for solutions (Corollary 4.7) and as a tool in the proof of Theorem B, and in the latter case as a tool to establish the uniqueness of the solution.

As an illustration of this state of affairs, consider again the family of effectivity functions E_n from Example 7.2. For $n \equiv 1 \pmod 3$ we have many implementable SCCs with effectivity function E_n (Theorem B), all of them between the nucleus (empty for some profiles) and the core. For $n \equiv -1 \pmod 3$ we have only one: the core (and nucleus) correspondence.

When the requirement of anonymity is dropped, effectivity functions for which the nucleus correspondence exists and is significantly different from the core correspondence do exist. We shall see this in the next section.

---

[13] We borrow this argument from Oren (1981, p. 217), where a statement of the required generalization may be found.

8 A Counterexample

The example that we shall construct in this section gives a negative answer to the question: Are the known necessary conditions for implementability also sufficient? In other words, we shall show that the situation described in Theorem B for the case of 3 alternatives does not carry over to the general case. Our example has 4 alternatives, and the same method may be used to construct similar examples with any number $\geqslant 4$ of alternatives.

Let A be a set of 4 alternatives and let $N = \{1, \ldots, 10\}$. Let $E = E(p, q)$ be the additive effectivity function (see Example 3.2) where p is the uniform probability measure on A and q is given by:

$$q(1) = \frac{2}{11}, \quad q(2) = \ldots = q(10) = \frac{1}{11}.$$

Clearly E is maximal. In fact, E may be considered as the proportional allocation of power for 10 individuals and 4 alternatives, corrected to obtain maximality[14]. Moreover, being an additive effectivity function, E is stable (by Moulin and Peleg 1982, Theorem 4.5). Translating (3.6) into blocking terminology and rounding up we obtain that for all $S \in \mathcal{P}(N)$ and all $B \in \mathcal{P}(A) - \{A\}$:

$$S \in \mathcal{B}(B) \Leftrightarrow 11q(S) \geqslant 3|B|, \tag{8.1}$$

where it is useful to think of $11q(S)$ as the number of members of S, with 1 counted twice. The minimal blocking coalitions for a single alternative are those with a "count" of 3. It is now clear that E obeys the SRB (Definition 7.1). Hence E is nucleus-stable by Theorem 7.3.

For a profile $R^N \in L^N$, define:

$$\text{Top}(R^N) = \{a \in A \mid \text{there exists } i \in N \text{ with } aR^i b \text{ for all } b \in A\},$$

$$H(R^N) = \text{Nu}(E, R^N) \cup [C(E, R^N) \cap \text{Top}(R^N)]. \tag{8.2}$$

---

[14] An interpretation of the correction is: 1, the chairman, is given a "double vote" to prevent deadlocks in the committee.

It is obvious that H is a SCC, it has the SPA property and is contained in the core correspondence of E. Yet, we shall show that H is not implementable.

Suppose that $G = (X^1, \ldots, X^{10}; \pi)$ implements H. Let $x^N \in X^N$, let $R^N \in L^N$ and let $a, b \in A$, $a \neq b$. The notation $\psi(x^N, R^N, a, b)$ will stand for a statement as follows:

$$\psi(x^N, R^N, a, b) \Leftrightarrow \text{for all } y^N \in X^N,$$

$$[y^i = x^i \text{ for all } i \in V_b^a(R^N)] \Rightarrow \pi(y^N) \neq b. \tag{8.3}$$

For given R^N, a and b, the truth of $\psi(x^N, R^N, a, b)$ depends only on the components of x^N corresponding to members of $V_b^a(R^N)$.

Now let $R^N \in L^N$ and $a \in A$. Then from the definition of implementation it follows that:

$$a \in H(R^N) \Leftrightarrow \text{there exists } x^N \in X^N \text{ such that for all } b \in A - \{a\}$$

$$\psi(x^N, R^N, a, b). \tag{8.4}$$

Let R^N and R_*^N be two profiles that satisfy (we denote the 4 alternatives by a, b, c and d):

$$V_b^a(R^N) = V_b^a(R_*^N) = \{1, 2, 3\},$$

$$V_c^a(R^N) = V_c^a(R_*^N) = \{1, 4, 5\},$$

$$V_d^a(R^N) = \{1, 6, 7, 8\},$$

$$V_d^a(R_*^N) = \{6, 7, 8\}.$$

By verifying the condition (4.6) in view of (8.1), one sees that $a \in C(E, R^N)$. Moreover, $a \in \text{Top}(R^N)$ since it is ranked first by 1. Therefore $a \in H(R^N)$. On the other hand, $a \notin H(R_*^N)$. Indeed, $a \notin \text{Nu}(E, R_*^N)$ since it is impossible to choose two disjoint coalitions, one contained in $\{1, 2, 3\}$ and the other in $\{1, 4, 5\}$, having each a "count" of at least 3; and $a \notin \text{Top}(R_*^N)$ since $\{1, 2, 3\} \cap \{1, 4, 5\} \cap \{6, 7, 8\} = \emptyset$.

We shall derive a contradiction using (8.4). As $a \in H(R^N)$, let $x^N \in X^N$ be as in the right hand side of (8.4). Suppose now that $x_*^N \in X^N$ agrees with x^N on $\{1, 2, 3, 4, 5\}$. Then $\psi(x_*^N, R_*^N, a, b)$ and $\psi(x_*^N, R_*^N, a, c)$ hold because $\psi(x^N, R^N, a, b)$ and $\psi(x^N, R^N, a, c)$ hold. Moreover, since $\{6, 7, 8\} \in \mathcal{B}(d)$ we may independently choose the components of x_*^N for $\{6, 7, 8\}$ so that $\psi(x_*^N, R_*^N, a, d)$ holds (by Moulin and Peleg

1982, Lemma 6.9, a coalition is α-effective for a set of alternatives with respect to H – in the sense of our Definition 3.3 – if and only if it is α-effective for that set with respect to G – in the sense that it can choose strategies for its members so that the outcome will be in that set irrespective of the others' strategies; here we use this for $\{6,7,8\}$ and the set $A-\{d\}$). So we can find $x_*^N \in X^N$ that will enable us to conclude from (8.4) that $a \in H(R_*^N)$ – a contradiction.

Here are a few words of explanation about this example. For the effectivity function chosen, the nucleus and the core correspondences both exist and are respectively the smallest and the largest implementable SCCs, and there is enough room between them to construct a SCC that has the SPA property but is not implementable. The key to this phenomenon lies in the fact that (8.4) imposes more on H in terms of inter-profile consistency than is captured by the SPA property. By generalizing the above argument, it is possible to obtain necessary conditions for implementability that go beyond the SPA property. However these conditions are complicated and less appealing, so we do not pursue this line.

Acknowledgements: I am grateful to Professors H. Moulin and B. Peleg for arousing my interest in the problem and for some valuable discussions. I am indebted to Professor C. d'Aspremont for several helpful remarks.

References

Dasgupta P, Hammond P, Maskin E (1979) The implementation of social choice rules: some general results on incentive compatibility. Review of Economic Studies 46:185–216

Maskin E (1979) Implementation and strong Nash-equilibrium. In: Laffont JJ (ed) Aggregation and revelation of preferences. North-Holland, Amsterdam, pp 433–439

Moulin H (1981) The proportional veto principle. Review of Economic Studies 48:407–416

Moulin H, Peleg B (1982) Cores of effectivity functions and implementation theory. Journal of Mathematical Economics 10:115–145

Oren I (1981) The structure of exactly strongly consistent social choice functions. Journal of Mathematical Economics 8:207–220

Peleg B (1982) Convex effectivity functions, R. M. 46. Center for Research in Mathematical Economics and Game Theory, The Hebrew University, Jerusalem

Received November 1985