Self-implementation of social choice correspondences in strong Equilibrium

Bezalel Peleg, Hans Peters

Self-implementation of social choice correspondences in strong equilibrium

RM/18/005

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O Box 616
NL- 6200 MD Maastricht
The Netherlands

Self-implementation of social choice correspondences in strong equilibrium

Bezalel Peleg[*] Hans Peters[†]

January 30, 2018

Abstract

A social choice correspondence is self-implementable in strong equilibrium if it is implementable in strong equilibrium by a social choice function selecting from the correspondence itself as a game form. We characterize all social choice correspondences implementable this way by an anonymous social choice function satisfying no veto power, given that the number of agents is large relative to the number of alternatives. It turns out that these are exactly the social choice correspondences resulting from feasible elimination procedures as introduced in Peleg (1978).

Keywords Implementation, strong equilibrium, social choice correspondence
JEL Classification C70, D71

1 Introduction

A social choice correspondence chooses alternatives based on the preferences of the agents. Generally speaking one looks for social choice correspondences with desirable properties, such as anonymity, Pareto optimality, and many more. The problem, as already studied in Hurwicz (1972), is that preferences may be private knowledge or, more generally, agents are entitled to report any preferences they wish, resulting in alternatives chosen on the basis of the wrong information, and thus in the desired properties of the social choice correspondence being violated. Requiring strategy-proofness of a social choice function, meaning that no agent can ever benefit from not reporting truthfully, is in general too strong and results in dictatorship (Gibbard, 1973, Satterthwaite, 1975).

Implementation theory is concerned with finding game forms (mechanisms, decentralized systems) of which the equilibrium (Nash, strong, etc.) alternatives in the game with the true preferences coincide with the alternatives assigned to those preferences by the social choice correspondence under consideration.

[*]The Federmann Center for the Study of Rationality and the Institute of Mathematics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel. Email: pelegba@math.huji.ac.il

[†]Department of Quantitative Economics, Maastricht University. Email: h.peters@maastrichtuniversity.nl

In particular since the work of Hurwicz (1972) there is a large literature on necessary and/or sufficient conditions for implementation of social choice correspondences under various equilibrium concepts, with Maskin (1999) as one of the basic contributions. For an overview of this literature up to the current millennium see Jackson (2001).

A well-recognized drawback of many of the game forms or mechanisms employed in implementation theory is that they tend to be fairly complicated and not easy to use in practice. For instance, they may require agents to report not just preferences but complete preference profiles, to report integer numbers, etc. In the present paper, we therefore ask what is still feasible by using what we call 'self-implementation': this means implementation by a game form that is simply a selection (social choice function) from the correspondence under consideration and, thus, requires the agents just to report their own preferences and nothing else. Apart from the simplicity of such a mechanism its use is also defendable in the sense that it is close to the social choice correspondence that is deemed desirable. Specifically, we ask the following question: which social choice correspondences are self-implementable in strong equilibrium (that is, strategy-profiles such that no coalition can gain by deviating, as introduced in Aumann, 1959)?

It turns out that under some natural additional conditions we are able to give a precise answer to this question: if the number of agents is not too small and the social choice function that selects from the correspondence and implements it is anonymous and satisfies 'no veto power' then the correspondence must result from a so-called feasible elimination procedure, as already introduced in Peleg (1978). The number of agents being not too small will be made precise and, together with the no veto power property boils down to this number being at least twice as large as the number of alternatives – a condition satisfied in most (political) elections. No veto power means that no agent on its own is able to exclude any alternative from being chosen – again a natural condition in larger elections. This result is quite involved and to some extent based on a selection from existing results in the literature; nevertheless, we present a completely self-contained proof.

The approach in Peleg (1978) and subsequent work (see Peleg and Peters, 2010) is different from implementation theory. Given a social choice correspondence one looks for a game form that exactly represents it, meaning that coalitions can achieve in the game form exactly what they can achieve in the social choice correspondence. Thus, this requirement is much stronger than in implementation theory. On the other hand, it is only required that for every (true) preference profile there exists at least one (Nash, strong, ...) equilibrium, and this is a much weaker condition than in implementation theory. One could say that the main result of this paper reveals a case where the approaches overlap.

Section 2 introduces the main concepts and Section 3 presents the main result. Most parts of the proof are shifted to the Appendix. Section 4 concludes.

Notations The following basic notations are used throughout. For a set D, $|D|$

denotes the cardinality of D, $P(D)$ the power set, i.e., the set of all subsets of D, and $P_0(D)$ the set of all nonempty subsets of D.

2 Self-implementation in strong equilibrium

Let A be the set of m *alternatives*, $m \geq 2$, and let $N = \{1, \ldots, n\}$, $n \geq 2$, be the set of *voters*. Subsets of N are called *coalitions*. Let L be the set of all *preferences*, i.e., complete, antisymmetric and transitive binary relations, on A. Then L^N is the set of all (preference) *profiles*. A *social choice correspondence* (SCC) is a function $H : L^N \to P_0(A)$. A *social choice function* (SCF) is a function $F : L^N \to A$. A social choice function F is a *selection* from a social choice correspondence H if $F(R^N) \in H(R^N)$ for every $R^N \in L^N$.

A *game form* is an $(n+1)$-tuple $g = (\Sigma^1, \ldots, \Sigma^n, \pi)$, where Σ^i is the *strategy set* of player (voter) $i \in N$, and $\pi : \Pi_{i=1}^n \Sigma^i \to A$ is the *outcome function*. For every $R^N \in L^N$ the pair (g, R^N) is a(n ordinal) game. A strategy profile $\sigma \in \Pi_{i=1}^n \Sigma^i$ is a *strong equilibrium* (Aumann, 1959) in the game (g, R^N) if there are no $S \in P_0(N)$ and $\widetilde{\sigma}^S \in \Pi_{i\in S}\Sigma^i$ such that $\pi(\widetilde{\sigma}^S, \sigma^{N\setminus S}) \neq \pi(\sigma)$ and $\pi(\widetilde{\sigma}^S, \sigma^{N\setminus S}) R^i \pi(\sigma)$ for all $i \in S$.[1]

A social choice correspondence H is *strong equilibrium implementable* if there is a game form $g = (\Sigma^1, \ldots, \Sigma^n, \pi)$ such that for every $R^N \in L^N$ we have

$$H(R^N) = \{\pi(\sigma) : \sigma \text{ is a strong equilibrium in } (g, R^N)\} \ .$$

In this case we also say that the game form g implements the SCC H in strong equilibrium.

A social choice function F can be identified with the game form in which the strategy set of each voter is the set L and the outcome function is F, i.e., to each strategy profile (preference profile) $Q^N \in L^N$ the outcome (alternative) $F(Q^N)$ is assigned. We denote this game form simply by F. Then (F, R^N) is a game for every $R^N \in L^N$.

Let H be a social choice correspondence. We call H *strong self-implementable* if there is a social choice function F such that

(i) $F(R^N) \in H(R^N)$ for every $R^N \in L^N$, and

(ii) $H(R^N) = \{F(Q^N) : Q^N$ is a strong equilibrium in $(F, R^N)\}$.

In words, the selection F from H implements H in strong equilibrium.

We assume that every SCC H (including every SCF, since this can be viewed as a single-valued SCC) occurring in the rest of the paper is *non-imposed*, i.e., for every $x \in A$ there is an $R^N \in L^N$ such that $H(R^N) = \{x\}$.

A well-known necessary condition (Maskin, 1999; see also Jackson, 2001) for H to be (self-)implementable is the following.

Maskin monotonicity For all $R^N = (R^1, \ldots, R^n)$, $Q^N = (Q^1, \ldots, Q^n) \in L^N$, and $x \in H(Q^N)$, if xQ^iy implies xR^iy for all $y \in A$ and $i \in N$, then $x \in H(R^N)$.

[1] Here, $\sigma^{N\setminus S}$ denotes the restriction of σ to $N \setminus S$. Similar notation will be used throughout the paper.

3 Main result

The main purpose of this section is to characterize all social choice correspondences H that are self-implementable in strong equilibrium if the number of voters is relatively large and the selection that implements H satisfies two natural properties, namely anonymity and no-veto power. The latter means that no voter on his own should be able to exclude any alternative from being chosen. We arrive at this theorem by combining a number of existing results in the literature, but our proof will be self-contained.

We start with the following concept, introduced by Peleg (1978). A social choice function F is *exactly and strongly consistent* (ESC) if for every $R^N \in L^N$ the game (F, R^N) has a strong equilibrium $Q^N \in L^N$ such that $F(Q^N) = F(R^N)$. We now immediately have the following result.

Lemma 3.1. *Let the selection F from the social choice correspondence H implement H in strong equilibrium. Then F is ESS.*

Proof. Let $R^N \in L^N$ and $x = F(R^N)$. Then $x \in H(R^N)$ and therefore there is a strong equilibrium Q^N of the game (F, R^N) such that $F(Q^N) = x$. Hence, $F(Q^N) = F(R^N)$. □

The SCCs of interest in this section are based on so-called feasible elimination procedures. Informally, first, assign weights $\beta(x) \in \mathbb{N}$ to the alternatives $x \in A$. Consider a preference profile and take an alternative x that is bottom ranked at least $\beta(x)$ times. Delete that alternative from the profile and at the same time delete $\beta(x)$ preferences where x is bottom ranked. Repeat this procedure until one alternative remains, which happens under appropriate conditions.

Formally, we have the following definition. A function $\beta : A \to \mathbb{N}$ such that $\sum_{x \in A} \beta(x) = n + 1$ will be called a *weight function.* Notice that if such a function exists then we must have $n + 1 \geq m$.

Definition 3.2. Let β be a weight function. Let $R^N \in L^N$. A *feasible elimination procedure* (f.e.p.) for R^N is a sequence $(x_1, C_1; \ldots; x_{m-1}, C_{m-1}; x_m)$ such that

(a) $A = \{x_1, \ldots, x_m\}$,

(b) $C_1, \ldots, C_{m-1}$ are pairwise disjoint subsets of N and $|C_j| = \beta(x_j)$ for all $j = 1, \ldots, m - 1$,

(c) $x_k R^i x_j$ for all $j = 1, \ldots, m - 1$, $k = j + 1, \ldots, m$, and $i \in C_j$.

Thus, in a feasible elimination procedure $(x_1, C_1; \ldots; x_{m-1}, C_{m-1}; x_m)$, by condition (c) alternative x_1 is bottom ranked for all voters in C_1 and by condition (b), $|C_1| = \beta(x_1)$. Now x_1 is deleted from R^N and also the preferences of the voters in C_1 are deleted. In the remaining profile, x_2 is bottom ranked for all voters in C_2 by condition (c), and by condition (b), $|C_2| = \beta(x_2)$, so that x_2 can be deleted and also the preferences of the voters in C_2. And so on and so forth. Observe that after deleting x_1 there are $n - \beta(x_1)$ voters left, after deleting x_2 there are $n-\beta(x_1)-\beta(x_2)$ voters left, and after deleting x_{m-1} there are $n-\beta(x_1)-\ldots-\beta(x_{m-1})=\beta(x_m)-1\geq 0$ voters left.

An important observation about f.e.p.'s is the following. Suppose an alternative x is bottom ranked by (at least) the voters in some coalition S with $|S|=\beta(x)$, in a profile $R^N \in L^N$. Then x must be eliminated in every f.e.p. for R^N. To see this suppose there is an f.e.p. in which x is not eliminated and let y be the alternative eliminated last, say via coalition T. Then the finally left voters form a coalition S' containing S. We have $\beta(y)+\beta(x)=|T|+|S'|+1$ by the foregoing, but also $|T|+|S'|\geq\beta(y)+\beta(x)$, a contradiction.

It is not difficult to see that there exists always at least one f.e.p. under the assumptions in the definition. If every alternative x_j is bottom ranked less than $\beta(x_j)$ times, then the total number of voters is at most $\sum_{j=1}^{m}\beta(x_j)-m$, which is equal to $n+1-m$ and therefore strictly smaller than n. A similar argument can be made after elimination of each alternative $x_1,\ldots,x_{m-2}$.

Let β be a weight function. An alternative x is R^N-*maximal* if there exists an f.e.p. $(x_1,C_1;\ldots;x_{m-1},C_{m-1};x)$. We denote

$$M_\beta(R^N)=\{x\in A : x \text{ is } R^N\text{-maximal}\}.$$

Thus, M_β depends on the exogenously chosen weights $\beta(x)\in\mathbb{N}$, $x\in A$, which can be varied as long as $\sum_{x\in A}\beta(x)=n+1$.

The following lemma repeats the known result that M_β is Maskin monotonic. For completeness, a proof can be found in the appendix. For a weight function β as in Definition 3.2 we use the notation $\beta(B)=\sum_{x\in B}\beta(x)$ for $B\subseteq A$.

Lemma 3.3. *Let β be a weight function. Then M_β is Maskin monotonic.*

Next, we provide a characterization of maximal alternatives.

Lemma 3.4. *Let β be a weight function. Let $x\in A$ and $R^N\in L^N$. The following statements are equivalent.*

(*i*) $x\in M_\beta(R^N)$.

(*ii*) *There are no $S\in P_0(N)$ and $B\in P_0(A)$ such that $|S|\geq\beta(A\setminus B)$, $x\in A\setminus B$, and $y\,R^i\,x$ for all $i\in S$ and $y\in B$.*

The following result says that M_β is self-implementable in strong equilibrium by any selection from it.

Proposition 3.5. *Let β be a weight function and let F be a selection from M_β. Then F implements M_β in strong equilibrium.*

Proof. (a) Let $R^N\in L^N$ and $x\in M_\beta(R^N)$. We show that there is a strong equilibrium Q^N of (F,R^N) such that $F(Q^N)=x$. Let $(x_1,C_1;\ldots;x_{m-1},C_{m-1};x)$ be an f.e.p. for R^N and consider the profile $Q^N\in L^N$ obtained from R^N by lowering x_j to the last position in the preferences of the voters in C_j, $j=1,\ldots,m-1$, leaving everything else in tact. Then $M_\beta(Q^N)=\{x\}$, hence $F(Q^N)=x$. Also, Q^N is a strong equilibrium of (F,R^N). Indeed assume on the contrary that there exist $S\in P_0(N)$ and $P^S\in L^S$ such that

$F(P^S, Q^{N\setminus S}) = z \neq x$ and zR^ix for all $i \in S$. Then $z = x_j$ for some $1 \leq j \leq m-1$. By the definition of an f.e.p., xR^iz for all $i \in C_j$, hence $S \cap C_j = \emptyset$. Since $|C_j| = \beta(z)$ and z is the last ranked alternative of Q^ℓ for all $\ell \in C_j$, we have that $z \notin M_\beta(P^S, Q^{N\setminus S})$, contradicting $F(P^S, Q^{N\setminus S}) = z$.

(b) Let Q^N be strong equilibrium of (F, R^N) with $F(Q^N) = x$. We show that $x \in M_\beta(R^N)$. It is sufficient to show that (ii) of Lemma 3.4 holds for x. Suppose not. Then there is an $S \in P_0(N)$ and $B \in P_0(A)$, $x \notin B$, such that yR^ix for all $y \in B$ and $i \in S$, and $|S| \geq \beta(A \setminus B)$. Consider a profile $P^S \in L^S$ with $A \setminus B$ at bottom for all voters in S. Then by the remarks following Definition 3.2, all elements of $A \setminus B$ will be eliminated in any f.e.p. for $(P^S, Q^{N\setminus S})$, so that $M_\beta(P^S, Q^{N\setminus S}) \subseteq B$, hence S has an improvement, a contradiction to the assumption that Q^N is strong equilibrium of (F, R^N). □

Before turning to a converse of Proposition 3.5 we introduce two additional possible properties of a social choice correspondence H. Of course, these properties also apply for a social choice function F, since a social choice function can be identified with a single-valued social choice correspondence.

Anonymity For all $R^N \in L^N$ and for all permutations π of N, $H(R^1, \ldots, R^n) = H(R^{\pi(1)}, \ldots, R^{\pi(n)})$.

No Veto Power For all $x \in A$ and $i \in N$, there is no $R^i \in L$ such that $x \notin H(R^i, R^{N\setminus\{i\}})$ for all $R^{N\setminus\{i\}} \in L^{N\setminus\{i\}}$.

Proposition 3.6. *Let social choice function F be ESC, anonymous, and satisfy No Veto Power, and let $n+1 \geq m$. Then there is a weight function β such that F is a selection from M_β.*

This proposition follows from earlier results in the literature, but for completeness we provide a self-contained proof in the appendix. The following theorem is a corollary to Propositions 3.5 and 3.6 and the main result of this section.

Theorem 3.7. *Let $n+1 \geq m$ and let the social choice function H be implementable in strong equilibrium by a selection F which is anonymous and satisfies No Veto Power. Then $H = M_\beta$ for some weight function β.*

Proof. By Lemma 3.1 and Proposition 3.6 it follows that there is a weight function β such that $F(R^N) \in M_\beta(R^N)$ for all $R^N \in L^N$. By Proposition 3.5, F implements M_β in strong equilibrium. Hence,

$$H(R^N) = \{F(Q^N) : Q^N \text{ is a strong equilibrium in } (F, R^N)\} = M_\beta(R^N)$$

for all $R^N \in L^N$, which completes the proof. □

Theorem 3.7 says, roughly, that if the number of voters is relatively large, then the only social choice correspondences which are self-implementable in a reasonable way in strong equilibrium are the correspondences M_β. Typically, in political elections the constraint $n+1 \geq m$ is satisfied and the conditions of

Anonymity and No Veto Power for a final selection of a candidate are natural if not compelling.

The conditions of Anonymity and No Veto Power in the theorem are on the selection F. It is not difficult to see that F can be anonymous but H not, or the other way around. If F satisfies No Veto Power then also H does, but the converse is not necessarily true. Since, thus, M_β in the theorem satisfies No Veto Power, it follows by the definition of an f.e.p. that $\beta(x) \geq 2$ for all $x \in A$.

4 Concluding remarks

Clearly, the approach in this paper leaves many open questions. We mention two of these. First, which social choice correspondences are self-implementable in strong equilibrium if the number of agents is relatively small – for instance, a small group of people in a restaurant has to make some common choices from a large menu of dishes? Second, what can be said about self-implementation in Nash equilibrium?

A Remaining proofs

A.1 Proofs of Lemmas 3.3 and 3.4

Proof of Lemma 3.3.[2] Let Q^N and R^N be as in the definition of Maskin monotonicity, and $x \in M_\beta(Q^N)$. Without loss of generality we assume that there is a voter v such that $Q^{N\setminus\{v\}} = R^{N\setminus\{v\}}$. Let $f^* = (x_1, C_1; \ldots; x_{m-1}, C_{m-1}; x)$ be an f.e.p. for Q^N, where $A = \{x_1, \ldots, x_{m-1}, x\}$. If $v \notin C_1 \cup \ldots \cup C_{m-1}$ then it is easy to see that f^* is still an f.e.p. for R^N, so that $x \in M_\beta(R^N)$. Now assume $v \in C_1 \cup \ldots \cup C_{m-1}$. If $v \in C_j$ with $j > 1$, then we may eliminate $x_1, \ldots, x_{j-1}$ and all voters in $C_1 \cup \ldots \cup C_{j-1}$ first, and next continue the argument with the remaining profile, where now all voters in C_j have x_j bottom ranked according to Q^{C_j}. So, without loss of generality, let $v \in C_1$.

The rest of the proof is based on a three step algorithm.

<u>Step 1</u> If the bottom alternative of R^v is equal to x_1, then f^* is still an f.e.p. for R^N and we are done. Otherwise, go to Step 2.

<u>Step 2</u> Let the bottom alternative of R^v be $x_\ell \neq x_1$, so $\ell \in \{2, \ldots, m-1\}$. If all voters in C_ℓ have x_ℓ as bottom alternative in R^N, then we can first eliminate x_ℓ via C_ℓ and go back to Step 1 for the reduced profile. Otherwise, go to Step 3.

<u>Step 3</u> Take $\hat{v} \in C_\ell$ with x_ℓ not as bottom alternative and note that the bottom alternative of $R^{\hat{v}} = Q^{\hat{v}}$ is some x_j with $j < \ell$ (since x_j must be eliminated before x_ℓ in f^*). Then modify C_ℓ to $\hat{C}_\ell = (C_\ell \cup \{v\}) \setminus \{\hat{v}\}$ and modify C_1 to $\hat{C}_1 = (C_1 \cup \{\hat{v}\}) \setminus \{v\}$. (In words, we switch v and $\hat{v}$.) Go back to Step 1.

[2] See Lemma 5.3.5 in Peleg (1984); or Remark 9.3.7 in Peleg and Peters (2010), based on Theorem 9.3.6 in the same source. In turn, the latter result goes back to Polishchuk (1978). More generally, Lemma 3.7 in Peleg and Peters (2017b) shows Maskin monotonicity of an extension of M_β.

Repeat this procedure until the final substitute of v in the modified C_1 has x_1 at bottom. Then we can apply an f.e.p. resulting in x, so that $x \in M_\beta(R^N)$. $\square$

Proof of Lemma 3.4.[3] For the implication $(i) \Rightarrow (ii)$, let $x \in M_\beta(R^N)$ and let $(x_1, C_1; \ldots; x_{m-1}, C_{m-1}; x)$ be an f.e.p. for R^N. Suppose there were S and B as in (ii). Write $B = \{x_{i_1}, \ldots, x_{i_{|B|}}\} \subseteq \{x_1, \ldots, x_{m-1}\}$, then $\left(\cup_{j=1}^{|B|} C_{i_j}\right) \cap S = \emptyset$ by definition of an f.e.p., and $|\cup_{j=1}^{|B|} C_{i_j}| = \beta(B)$. Hence $|S| + |\cup_{j=1}^{|B|} C_{i_j}| \geq \beta(A \setminus B) + \beta(B) = n + 1$, a contradiction.

We prove the implication $(ii) \Rightarrow (i)$ by induction on the number of alternatives m. Let $x \in A$ and assume that (ii) holds.

If $m = 2$, say $A = \{x, y\}$, then there is no $S \in P_0(N)$ such that $|S| \geq \beta(x)$ and yR^ix for all $i \in S$, so that $M_\beta(R^N) = \{x\}$.

Now suppose that $m > 2$ and that the implication $(ii) \Rightarrow (i)$ holds if there are less than m alternatives. For every $B \in P_0(A \setminus \{x\})$ denote $S_B = \{i \in N : yR^ix \text{ for all } y \in B\}$. Then (ii) is equivalent to

$$|S_B| < \beta(A \setminus B) \text{ for all } B \in P_0(A \setminus \{x\}) \tag{1}$$

hence to

$$|N \setminus S_B| \geq \beta(B) \text{ for all } B \in P_0(A \setminus \{x\}). \tag{2}$$

We consider two cases.

Case 1 There exists $\widetilde{B} \in P_0(A \setminus \{x\})$ with $|\widetilde{B}| \leq m - 2$ and $|N \setminus S_{\widetilde{B}}| = \beta(\widetilde{B})$.

For this case we consider the two following subproblems:

- $N_1 = N \setminus S_{\widetilde{B}}$, $A_1 = \widetilde{B} \cup \{x\}$, $\beta_1(y) = \beta(y)$ for all $y \in \widetilde{B}$, $\beta_1(x) = 1$, and $R_1^i = R^i_{|A_1}$ for all $i \in N_1$.[4]

- $N_2 = S_{\widetilde{B}}$, $A_2 = A \setminus \widetilde{B}$, $\beta_2(y) = \beta(y)$ for all $y \in A_2$, and $R_2^i = R^i_{|A_2}$ for all $i \in N_2$.

We next show that (1) holds for the first subproblem. If not, then there is a $B \in P_0(\widetilde{B})$ such that $|T| \geq \beta_1(A_1 \setminus B)$, where $T = \{i \in N_1 : yR_1^ix \text{ for all } y \in B\}$. Then $|T \cup S_{\widetilde{B}}| = |T| + |S_{\widetilde{B}}| \geq [\beta_1(x) + \beta(\widetilde{B}) - \beta(B)] + [n - \beta(\widetilde{B})] = \beta(A \setminus B)$, hence $|S_B| \geq \beta(A \setminus B)$, which is a violation of (1) for the original problem. Therefore, (1) must hold for the first subproblem, implying that $x \in M_{\beta_1}(R_1^{N_1})$ by induction.

Similarly, suppose that (1) does not hold for the second subproblem. Then there is a $B \in P_0(A \setminus (\widetilde{B} \cup \{x\}))$ such that $|T| \geq \beta_2(A_2 \setminus B)$, where now $T = \{i \in S_{\widetilde{B}} : yR_2^ix \text{ for all } y \in B\}$. Then $|T \cup (N \setminus S_{\widetilde{B}})| = |T| + |N \setminus S_{\widetilde{B}}| \geq [\beta(A) - \beta(B) - \beta(\widetilde{B})] + \beta(\widetilde{B}) = \beta(A \setminus B)$, which is a violation of (1) for the original problem. We conclude that (1) must hold for the second subproblem as well, so that $x \in M_{\beta_2}(R_2^{N_2})$ by induction.

[3] Also this result can be deduced from Theorem 9.3.6 in Peleg and Peters (2010). It is included as Lemma 3.5 in Peleg and Peters (2017a).

[4] $R^i_{|B}$ denotes the restriction of R^i to B.

Now let $(z_1, C_1; \ldots; z_{|\widetilde{B}|}, C_{|\widetilde{B}|}; x)$ be an f.e.p. for the first subproblem and let $(u_1, D_1; \ldots; u_{m-1-|\widetilde{B}|}, D_{m-1-|\widetilde{B}|}; x)$ be an f.e.p. for the second subproblem. Since, in particular, $yR^i x$ for all $y \in \widetilde{B}$ and $i \in N_2 = S_{\widetilde{B}}$, it follows that

$$(u_1, D_1; \ldots; u_{m-1-|\widetilde{B}|}, D_{m-1-|\widetilde{B}|}; z_1, C_1; \ldots; z_{|\widetilde{B}|}, C_{|\widetilde{B}|}; x)$$

is an f.e.p. for the original problem, implying that in this case we have $x \in M_\beta(R^N)$.

Case 2 For all $\widetilde{B} \in P_0(A \setminus \{x\})$ with $|\widetilde{B}| \leq m-2$ we have $|N \setminus S_{\widetilde{B}}| > \beta(\widetilde{B})$.

Suppose there is an $\ell \in N$ such that x is not ranked at the last or second last position in R^ℓ, and let $\widehat{y}$ be the alternative ranked right below x. We switch x and $\widehat{y}$ in voter ℓ's preference to obtain a new preference $\widehat{R}^\ell$ and a new preference profile $\widehat{R}^N = (R^1, \ldots, R^{\ell-1}, \widehat{R}^\ell, R^{\ell+1}, \ldots, R^N)$ that still satisfies (2): for any set B with $|B| \leq m-2$ this holds because of the strict inequality in Case 2, and for $B = A \setminus \{x\}$ this holds since x is not ranked last in $\widehat{R}^\ell$.

If Case 1 applies to $\widehat{R}^N$, then $x \in M_\beta(\widehat{R}^N)$. Thus, by Lemma 3.3, $x \in M_\beta(R^N)$. If Case 1 does not apply to $\widehat{R}^N$, then we repeat this step for some voter $\ell' \in N$ with x not ranked last or second last at $\widehat{R}^{\ell'}$, and so on, until either Case 1 applies or there is no voter left with x not ranked at the last or second last position.

In the latter case, we have a profile, say $\widetilde{R}^N$, for which still (2) holds and with x ranked last or second last for each voter $i \in N$. Observe that y is last ranked for all voters in $N \setminus S_{\{y\}}$ for all $y \in A \setminus \{x\}$. Also, by (2), $|N \setminus S_{\{y\}}| \geq \beta(y)$ for all $y \in A \setminus \{x\}$. It follows that in any f.e.p. for $\widetilde{R}^N$ every $y \in A \setminus \{x\}$ is bottom ranked by at least $\beta(y)$ voters and therefore eliminated, so that $M_\beta(\widetilde{R}^N) = \{x\}$. By Lemma 3.3 again, $x \in M(R^N)$.

By (2), Cases 1 and 2 are exhaustive, which completes the proof of the lemma. □

A.2 Proof of Proposition 3.6

We now turn to the proof of Proposition 3.6. It will be convenient to introduce some terminology related to effectivity functions.[5] Let F be a social choice function and let $S \subseteq N$ and $B \subseteq A$. Then S is (F-)*effective for* B is there is $R^S \in L^S$ such that $F(R^S, Q^{N\setminus S}) \in B$ for all $Q^{N\setminus S} \in L^{N\setminus S}$. For every $x \in A$ define the integer $b(x)$ (the 'blocking coefficient' of x) by

$$b(x) = \min\{|S| : \ S \subseteq N \text{ is effective for } A \setminus \{x\}\ \} .$$

By non-imposition of F, we have $1 \leq b(x) \leq n$ for all $x \in A$. We write $b(B)$ for $\sum_{x \in B} b(x)$, $B \subseteq A$. Of course, $b(\cdot)$ depends on F but this will be suppressed from notation if confusion is unlikely.

[5]These functions have been first formally introduced in Moulin and Peleg (1982). Here we just use some of the associated terminology.

We start with three useful observations.[6]

Lemma A.1. *Let the SCF F be anonymous. Let $S \subseteq N$ and $B \subseteq A$ such that $|S| \geq b(A \setminus B)$. Then S is effective for B.*

Proof. Write $A \setminus B = \{x_1, \ldots, x_k\}$, where $k \geq 0$. Let $S_1, \ldots, S_k$ be a partition of S such that $|S_j| \geq b(x_j)$ for each $j = 1, \ldots, k$, and let $R^{S_j} \in L^{S_j}$ such that $F(R^{S_j}, Q^{N \setminus S_j}) \in A \setminus \{x_j\}$ for each $j = 1, \ldots, k$ and $Q^{N \setminus S_j} \in L^{N \setminus S_j}$. Then $F(R^S, Q^{N \setminus S}) \in B$ for all $Q^{N \setminus S} \in L^{N \setminus S}$. So S is effective for B. □

Lemma A.2. *Let the SCF F be ESC and let $S \subseteq N$ be effective for $B \subseteq A$. Let $R^N \in L^N$ and $x \in A \setminus B$ such that $yR^i x$ for all $y \in B$ and $i \in S$. Then $F(R^N) \neq x$.*

Proof. Suppose on the contrary that $F(R^N) = x$ and let Q^N be a strong equilibrium in (F, R^N) with $F(Q^N) = x$. Since S is effective for B, there is $P^S \in L^S$ such that $F(P^S, Q^{N \setminus S}) \in B$, contradicting that Q^N is a strong equilibrium in (F, R^N). □

Lemma A.3. *Let the SCF F be ESC and anonymous, and assume that $b(A) = n + 1$. Then F is a selection from M_b.*

Proof. Let $R^N \in L^N$ and $x = F(R^N)$. Let $B \subseteq A$, $S \subseteq N$, $|S| \geq b(A \setminus B)$, and $x \in A \setminus B$. In order to prove that $x \in M_b(R^N)$, it is by Lemma 3.4 sufficient to prove that we do not have $yR^i x$ for all $y \in B$ and $i \in S$.

On the contrary, suppose that $yR^i x$ for all $y \in B$ and $i \in S$. Since $|S| \geq b(A \setminus B)$, Lemma A.1 implies that S is effective for B. Then Lemma A.2 implies that $F(R^N) \neq x$, a contradiction. □

Notice that in order to obtain Proposition 3.6 we may try and derive the condition $b(A) = n + 1$ in Lemma A.3. This is, essentially, what is done in the remainder of the proof.

Lemma A.4. *Let the SCF F be ESC, $S \subseteq N$, $B \subseteq A$, and suppose that for every $Q^{N \setminus S} \in L^{N \setminus S}$ there is $P^S \in L^S$ such that $F(P^S, Q^{N \setminus S}) \in B$. Then S is effective for B.*[7]

Proof. On the contrary, suppose that for every $Q^S \in L^S$ there is $P^{N \setminus S} \in L^{N \setminus S}$ such that $F(Q^S, P^{N \setminus S}) \in A \setminus B$. Consider a profile $R^N \in L^N$ such that $xR^i y$ and $yR^j x$ for every $i \in S$, $j \in N \setminus S$, $x \in B$, and $y \in A \setminus B$. Let $z = F(R^N)$ and let Q^N be a strong equilibrium of (F, R^N) with $F(Q^N) = z$. If $z \in A \setminus B$ then S can improve by a profile P^S as in the statement of the lemma. If $z \in B$ then $N \setminus S$ can improve by a profile $P^{N \setminus S}$ as above. □

In what follows we will use the notion of a generalized partition or *g-partition* of a set, which is a partition in which some elements may be empty.

[6] Many of the arguments in this part are based on Chapter 10 in Peleg and Peters (2010) and the references therein.

[7] This lemma states that the effectivity function associated with F is 'maximal'. See Moulin and Peleg (1982) or Peleg (1984).

Lemma A.5. *Let the SCF F be ESC. Then there are no $p \geq 2$, partition $B_1, \ldots, B_p$ of A and g-partition $S_1, \ldots, S_p$ of N such that $N \setminus S_i$ is effective for B_i, for every $i = 1, \ldots, p$.*

Proof. Suppose not, so (g-)partitions as in the lemma exist. Consider a profile R^N as in the following table:

$$\begin{array}{cccc} S_1 & S_2 & \cdots & S_p \\ \hline B_2 & B_3 & \cdots & B_1 \\ \vdots & \vdots & & \vdots \\ B_p & B_1 & \cdots & B_{p-1} \\ B_1 & B_2 & \cdots & B_p \end{array}$$

(meaning that every member of coalition S_1 prefers all alternatives of B_2 over all alternatives of B_3, all alternatives of B_3 over all alternatives of B_4, and so on and so forth). Now by Lemma A.2, $F(R^N) \notin B_i$ for every $i = 1, \ldots, p$. Since $\cup_{i=1}^{p} B_i = A$, this is a contradiction. □

Lemma A.6. *Let the SCF F be ESC and satisfy NVP. Then there are no partition $\{x\}, B_1, B_2$ of A and g-partition S, T_1, T_2 of N such that $|S| = b(x)$ and $N \setminus T_j$ is effective for B_j for $j = 1, 2$.*

Proof. Suppose not, so (g-)partitions as in the lemma exist.

First, suppose $S = N$. Then for every $i \in N$, $|N \setminus \{i\}| < |S| = b(x)$. Therefore, for every $Q^{N\setminus\{i\}} \in L^{N\setminus\{i\}}$ there is $P^i \in L$ such that $F(P^i, Q^{N\setminus\{i\}}) = x$, so that by Lemma A.4, $\{i\}$ is effective for x. Since $|A| \geq 2$ this violates NVP of F. Thus, $S \neq N$ and $b(x) < n$. By NVP, also $b(x) > 1$. So $|S| \geq 2$ and $T_1 \cup T_2 \neq \emptyset$.

Let now S_1, S_2 be a partition of S and consider a profile R^N as in the following table:

$$\begin{array}{cccc} S_1 & S_2 & T_1 & T_2 \\ \hline B_2 & B_1 & \{x\} & \{x\} \\ B_1 & B_2 & B_2 & B_1 \\ \{x\} & \{x\} & B_1 & B_2 \end{array}$$

Since $S = S_1 \cup S_2$ is effective for $A \setminus \{x\} = B_1 \cup B_2$ we have by Lemma A.2 that $F(R^N) \neq x$. Without loss of generality we assume that $F(R^N) \in B_1$. Let Q^N be a strong equilibrium in (F, R^N) with $F(Q^N) = F(R^N)$, hence $F(Q^N) \neq x$.

<u>Case 1</u>: xQ^iy for some $i \in S$, without loss of generality $i \in S_1$, and $y \in A \setminus \{x\}$.

In this case consider the partition $\{x\}, \{y\}, A \setminus \{x, y\}$ of A and the g-partition $S \setminus \{i\}, \{i\}, T_1 \cup T_2$ of N. Since $|S \setminus \{i\}| < b(x)$ we have that $N \setminus (S \setminus \{i\})$ is effective for $\{x\}$ by Lemma A.4. By NVP and Lemma A.4, $N \setminus \{i\}$ is effective for $\{y\}$. Hence, by Lemma A.5, $N \setminus (T_1 \cup T_2)$ is not effective for $A \setminus \{x, y\}$. In turn, again by Lemma A.4, this implies that $T_1 \cup T_2$ is effective for $\{x, y\}$. Consider a profile $P^{T_1 \cup T_2} \in L^{T_1 \cup T_2}$ such that xP^jyP^jz for all $j \in T_1 \cup T_2$ and $z \in A \setminus \{x, y\}$. Then by Lemma A.2, $F(P^{T_1 \cup T_2}, Q^S) \in \{x, y\}$. Since xQ^iy and since $T_1 \cup T_2 \cup \{i\} = N \setminus (S \setminus \{i\})$ is effective for $\{x\}$, again by Lemma A.2,

$F(P^{T_1 \cup T_2}, Q^S) \neq y$. Hence, $F(P^{T_1 \cup T_2}, Q^S) = x$. This contradicts that Q^N is a strong equilibrium in (F, R^N).

Case 2: $yQ^i x$ for all $i \in S$ and $y \in A \setminus \{x\}$.

In this case, consider the partition $\{x\}, B_1, B_2$ of A and the g-partition $S_2, S_1 \cup T_1, T_2$ of N. Since $|S_2| < b(x)$ we have by Lemma A.4 that $N \setminus S_2$ is effective for $\{x\}$. By assumption, $N \setminus T_2$ is effective for B_2. Hence by Lemma A.5, $N \setminus (S_1 \cup T_1)$ is not effective for B_1, which in turn by Lemma A.4 implies that $S_1 \cup T_1$ is effective for $A \setminus B_1$. Consider a profile $P^{S_1 \cup T_1} \in L^{S_1 \cup T_1}$ such that $yP^j xP^j z$ for all $j \in S_1 \cup T_1$, $y \in B_2$, and $z \in B_1$. By Lemma A.2, $F(P^{S_1 \cup T_1}, Q^{S_2 \cup T_2}) \notin B_1$. Since by assumption $S_1 \cup S_2 \cup T_1$ is effective for B_2, by Case 2 $yQ^i x$ for all $y \in B_2$ and $i \in S$, and $N \setminus T_2$ is effective for B_2, we have by Lemma A.2 that $F(P^{S_1 \cup T_1}, Q^{S_2 \cup T_2}) \neq x$. Hence $F(P^{S_1 \cup T_1}, Q^{S_2 \cup T_2}) \in B_2$. Since $F(Q^N) = F(R^N) \in B_1$, $S_1 \cup T_1$ has an improvement, contradicting that Q^N is a strong equilibrium of (F, R^N). □

Lemma A.7. *Let the SCF F be ESC and satisfy NVP, and $1 \leq k \leq m - 2$. Then there are no partition $\{x_1\}, \ldots, \{x_k\}, B_1, B_2$ of A and g-partition $S_1, \ldots, S_k, T_1, T_2$ of N such that $|S_i| = b(x_i)$ for each $i = 1, \ldots, k$, $N \setminus T_1$ is effective for B_1, and $N \setminus T_2$ is effective for B_2.*

Proof. The proof is by induction on k. For $k = 1$ this is Lemma A.6. Let $2 \leq k \leq m - 2$ and suppose that the statement in the lemma holds for $k - 1$. Suppose, on the contrary, that the statement does not hold for k, and let $\{x_1\}, \ldots, \{x_k\}, B_1, B_2$ and $S_1, \ldots, S_k, T_1, T_2$ be as in the lemma. Since $S_i \neq \emptyset$ for every $i = 1, \ldots, k$, we have $\emptyset \neq S_k \cup T_1 \neq N$. By Lemma A.4, either $S_k \cup T_1$ is effective for $A \setminus (\{x_k\} \cup B_1)$ or $N \setminus (S_k \cup T_1)$ is effective for $\{x_k\} \cup B_1$. In the first case, Lemma A.6 is violated for the partition $\{x_k\}, B_1, A \setminus (\{x_k\} \cup B_1)$ of A and the g-partition $S_k, T_1, N \setminus (S_k \cup T_1)$ of N. In the second case, the induction hypothesis is violated for the partition $\{x_1\}, \ldots, \{x_{k-1}\}, \{x_k\} \cup B_1, B_2$ of A and the g-partition $S_1, \ldots, S_{k-1}, S_k \cup T_1, T_2$ of N. □

The next lemma says that an ESC social choice function is 'subadditive'.[8]

Lemma A.8. *Let the SCF F be ESC, let $S_1 \subseteq N$ be effective for $B_1 \subseteq A$ and let $S_2 \subseteq N$ be effective for $B_2 \subseteq A$, such that $B_1 \cap B_2 = \emptyset$. Then $S_1 \cap S_2$ is effective for $B_1 \cup B_2$.*

Proof. (a) Say that coalition S is *s-effective* for a set of alternatives B if there is a partition $B_1, \ldots, B_k$ of B and there are coalitions $S_1, \ldots, S_k$ such that S_j is effective for B_j, $j = 1, \ldots, k$, and $S = \cap_{j=1}^k S_j$. Clearly, if S is effective for B then S is also s-effective for B by taking $k = 1$, $S_1 = S$, $B_1 = B$. We will prove the converse, which will imply the lemma.

(b) We first prove that if S is s-effective for B, then $N \setminus S$ is not s-effective for $A \setminus B$. Suppose the latter were not the case, i.e., both S is s-effective for B and $N \setminus S$ is not s-effective for $A \setminus B$. Let $B_1, \ldots, B_k$ and $C_1, \ldots, C_\ell$ be the associated partitions of B and $A \setminus B$, and let $S_1, \ldots, S_k$ and $T_1, \ldots, T_\ell$

[8]Cf. Moulin (1983).

be the associated coalitions, hence $S = \cap_{j=1}^{k} S_j$ and $N \setminus S = \cap_{h=1}^{\ell} T_h$. List $S_1, \ldots, S_k, T_1, \ldots, T_\ell$ as $V_1, \ldots, V_p$ and list the associated sets of alternatives as $D_1, \ldots, D_p$ (where $p = k + \ell$). Then for every $i \in N$ there is $q \in \{1, \ldots, p\}$ such that $i \notin V_q$. Consider a preference profile R^N such that for every $i \in N$, $D_{q+1} R^i D_{q+2} R^i \ldots R^i D_p R^i D_1 R^i \ldots R^i D_q$. Let $x \in A$. If $x \in D_q$ for some $q > 1$ then $D_{q-1} R^i x$ for all $i \in V_{q-1}$, so that by Lemma A.2 we have $F(R^N) \neq x$. If $x \in D_1$ then $D_p R^i x$ for all $i \in V_p$, so that again by Lemma A.2 we have $F(R^N) \neq x$. This is not possible, hence we have that if S is s-effective for B, then $N \setminus S$ is not s-effective for $A \setminus B$.

(c) Now, finally, assume that S is s-effective for B. Then by part (b), $N \setminus S$ is not s-effective for $A \setminus B$, hence by part (a), $N \setminus S$ is not effective for $A \setminus B$. By Lemma A.4, S is effective for B. This concludes the proof of the lemma. □

The final lemma we need is the following.

Lemma A.9. *Let the SCF F be ESC and satisfy NVP. Let $0 \leq k \leq m - 2$. Then there are no partition $\{x_1\}, \ldots, \{x_m\}$ of A and g-partition $S_1, \ldots, S_m$ of N such that $|S_j| = b(x_j)$ for $j = 1, \ldots, k$ and $|N \setminus S_j|$ is effective for $\{x_j\}$ for $j = k + 1, \ldots, m$.*

Proof. For $k = 0$ this follows from Lemma A.5. Now let $k > 0$. Suppose on the contrary that we had $\{x_1\}, \ldots, \{x_m\}$ and $S_1, \ldots, S_m$ as in the lemma. By repeated application of Lemma A.8 we have that $N \setminus (S_{k+1} \cup \ldots \cup S_{m-1})$ is effective for $\{x_{k+1}, \ldots, x_{m-1}\}$. Now the partition $\{x_1\}, \ldots, \{x_k\}, \{x_{k+1}, \ldots, x_{m-1}\}, \{x_m\}$ and g-partition $S_1, \ldots, S_k, T_1, T_2$ with $T_1 = S_{k+1} \cup \ldots \cup S_{m-1}$ and $T_2 = S_m$ violate Lemma A.7. □

Proof of Proposition 3.6. In view of Lemma A.3, it is sufficient to prove that $b(A) = n + 1$. Clearly, $b(A) \geq n + 1$, otherwise N would have some profile R^N such that $F(R^N) \notin A$, which is clearly impossible. Write $A = \{x_1, \ldots, x_m\}$. We distinguish two cases.

Case 1 $b(A) \geq n + m$. Then $n \leq b(A) - m = \sum_{j=1}^{m} (b(x_j) - 1)$, so that there is a g-partition $S_1, \ldots, S_m$ of N with $|S_j| \leq b(x_j) - 1$ for every $j = 1, \ldots, m$, which by using Lemma A.4 violates Lemma A.9 for $k = 0$.

Case 2 $b(A) = n + (m - k)$ for some $k \in \{1, \ldots, m - 2\}$. In this case, let S_j, $j = 1, \ldots, k$, be coalitions with $|S_j| = b(x_j)$. Since

$$\begin{aligned}
\sum_{j=1}^{k} |S_j| &= b(A) - (b(x_{k+1} + \ldots + b(x_m)) \\
&= n + (m - k) - (b(x_{k+1} + \ldots + b(x_m)) \\
&\leq n + (m - k) - (m - k) \\
&= n
\end{aligned}$$

the S_j can be chosen disjoint. Also,

$$\begin{aligned} n - \sum_{j=1}^{k} |S_j| &= n - (b(A) - \sum_{j=k+1}^{m} b(x_j)) \\ &= n - n - (m-k) + \sum_{j=k+1}^{m} b(x_j)) \\ &= \sum_{j=k+1}^{m} (b(x_j) - 1) \end{aligned}$$

so that we can find disjoint $S_{k+1}, \ldots, S_m$ with $|S_j| = b(x_j) - 1$ for all $j = k+1, \ldots, m$, hence, by Lemma A.4, $N \setminus S_j$ is effective for $\{x_j\}$. This is again a violation of Lemma A.9.

Thus, $b(A) = n + 1$, which concludes the proof. □

References

Aumann RJ (1959) Acceptable points in general cooperative n-person games. In: Contributions to the theory of games, Vol. IV, Annals of mathematic studies no. 40. Princeton University Press, Princeton, NJ

Gibbard A (1973) Manipulation of voting schemes: A general result. Econometrica 41:587–602

Hurwicz L (1972) On informationally decentralized systems. In Mcguire CB, Radner R: Decision and Organization. North-Holland, Amsterdam

Jackson MO (2001) A crash course in implementation theory. Social Choice and Welfare 18:655–708

Maskin E (1999) Nash equilibrium and welfare optimality. The Review of Economic Studies 66:23–38

Moulin H (1983) The strategy of social choice. North-Holland, Amsterdam

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

Peleg B (1978) Consistent voting systems. Econometrica 46:153–161

Peleg B (1984) Game theoretic analysis of voting in committees. Cambridge University Press, Cambridge UK

Peleg B, Peters H (2010) Strategic social choice: stable representations of constitutions. Springer, Heidelberg

Peleg B, Peters H (2017a) Feasible elimination procedures in social choice: an axiomatic characterization. Research in Economics 71:43–50

Peleg B, Peters H (2017b) Choosing k from m: feasible elimination procedures revisited. Games and Economic Behavior 103:254–261

Polishchuk (1978) Monotonicity and uniqueness of consistent voting systems. Center for Research in Mathematical Economics and Game Theory, Hebrew University of Jerusalem

Satterthwaite MA (1975) Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. Journal of Economic Theory 10:187–207