Representation of constitutions under incomplete information

Author(s): Bezalel Peleg and Shmuel Zamir

Source: *Economic Theory*, Vol. 57, No. 2 (October 2014), pp. 279-302
Published by: Springer
Stable URL: https://www.jstor.org/stable/43562896

Econ Theory (2014) 57:279–302
DOI 10.1007/s00199-014-0816-0

RESEARCH ARTICLE

Representation of constitutions under incomplete information

Bezalel Peleg and Shmuel Zamir

Received: 4 November 2013 / Accepted: 17 April 2014 / Published online: 1 May 2014

Abstract We model constitutions by effectivity functions. We assume that the constitution is common knowledge among the members of the society. However, the preferences of the citizens are private information. We investigate whether there exist decision schemes (i.e., functions that map profiles of (dichotomous) preferences on the set of outcomes to lotteries on the set of social states), with the following properties: (i) The distribution of power induced by the decision scheme is identical to the effectivity function under consideration; and (ii) the (incomplete information) game associated with the decision scheme has a Bayesian Nash equilibrium in pure strategies. If the effectivity function is monotonic and superadditive, then we find a class of decision schemes with the foregoing properties.

Keywords Effectivity function · Incomplete information · Decision scheme · Bayesian Nash equilibrium

JEL Classification C62 · C70 · D82

Introduction

Following Gardenfors (1981) and Peleg (1998), we model constitutions by effectivity functions. Formally, an *effectivity function* is the coalitional function of a game form, that is, a coalitional game form [see Abdou and Keiding (1991, p. 28)]. We assume that

B. Peleg (✉) · S. Zamir
Center for the Study of Rationality, The Hebrew University,
Givat-Ram Campus, 91904 Jerusalem, Israel
e-mail: pelegba@math.huji.ac.il

S. Zamir
e-mail: zamir@math.huji.ac.il

the constitution is common knowledge among the members of the society. However, the preferences of the members of the society over the set of social states are private information. In order to enable the citizens to exercise their rights and comply with their obligations according to the constitution, we need some kind of a game form or mechanism to represent it (see Peleg 1998 and Peleg and Peters 2010). In this paper, we represent constitutions by *decision schemes*, that is, functions from profiles of preferences of citizens to probability distributions over the set of social states. A decision scheme is a representation of a constitution if the power distribution (among coalitions of players) induced by it coincides with the constitution. A representation induces a Bayesian game whose players are the members of the society as we shall see in Sect. 2. We shall investigate various kinds of representations for which the induced incomplete information game possesses a pure Bayesian Nash equilibrium.

The following simple example may help the reader to become familiar with the foregoing concepts. It is a modification of a famous example of Gibbard (1974).

Example 1 Let $N = \{1, 2\}$ be a society. Assume that each member has two shirts, one white and one blue, and he must wear exactly one of them. Then there are four social states: ww, wb, bw, and bb, where ww means that they both wear white shirts etc. Assume further that each citizen can freely choose the color of his shirt. Then the constitution, that is the associated effectivity function E, is given by: $E(\{1\}) = \{\{ww, wb\}^+, \{bw, bb\}^+\}$; $E(\{2\}) = \{\{ww, bw\}^+, \{wb, bb\}^+\}$; and $E(N)$ is the set of all nonempty subsets of A, where $A = \{ww, wb, bw, bb\}$ is the set of all social states and for any $B \subseteq A$, we denote by B^+ the collection of all supersets of B. Assume now that player 1 has two types 1_w and 1_b, and player 2 believes that they have the same probability. Let W be the set of all complete and transitive (weak) orderings of A. A decision scheme is a function $d : W^N \to \Delta(A)$ where $\Delta(A)$ is the set of all probability distributions on A. In Sect. 1.1, we shall find a representation for E. To complete our example, we need to specify von-Neumann Morgenstern utility functions for 1_w, 1_b, and 2. We shall do this in Sect. 2 and then compute a pure Bayesian Nash equilibrium for the induced Bayesian game.

From a broad perspective, our study belongs to the vast literature on the tension between social welfare and the distribution of rights starting with Arrow (1951) (see Suzumura 2011, for a comprehensive survey of this field). On the one hand, our approach allows for any reasonable distribution of group rights. Thus, we can avoid dictatorial decision schemes. On the other hand, because we insist on precise representation of group rights, we might lose incentive compatibility of some of our Bayesian Nash equilibria. (We prove existence of Bayesian Nash equilibria in pure strategies.) Thus, although our representing decision schemes are ex-post Pareto optimal, we might only obtain Pareto optimality with respect to reported preferences. Hence, we do not fully avoid Sen's Paradox of the Paretian liberal. Comparing with the results for the case of complete information where there is always at least one Pareto optimal Nash equilibrium (for each profile of preferences), we conclude that there is a price to pay for generalizing the model of Peleg (1998) and Peleg et al. (2002) to allow incomplete information, namely we might lose Pareto optimality of the resulting social state.

Our work is not the first one that investigates representations of power structures under incomplete information; d'Aspremont and Peleg (1988) studies representations of simple games by decision schemes under the same assumptions of incomplete information. (A simple game is an example of a neutral effectivity function.) There are two significant differences between the two papers: (i) d'Aspremont and Peleg use a slightly stronger notion of representation; and (ii) they focus on a stronger notion of equilibrium, namely ordinal Bayesian (incentive compatible) equilibria.

We now review briefly the contents of the paper. Our model and some basic definitions are introduced in the first half of Sect. 1. The rest of Sects. 1 and 1.1 is devoted to recalling some results on the uniform core (due to Abdou and Keiding 1991). The uniform core of an effectivity function plays an important role throughout our paper. In Sect. 2, we consider a society whose members have incomplete information on each other's preferences. We prove, under mild restrictions, that the constitution of the society can be represented by a decision scheme such that the resulting game (of incomplete information) has a Bayesian Nash equilibrium in pure strategies. We further show that the decision scheme may be chosen to be ex-post Pareto optimal and that we may restrict ourselves to dichotomous preferences.

1 The model

- Let $N = \{1, 2, \ldots, n\}$ be the set of *players* (voters).
- Let $A = \{a_1, a_2, \ldots, a_m\}$ be the set of *alternatives*, $m \geq 2$.
- For a finite set D let $P(D) = \{D' | D' \subseteq D\}$ and $P_0(D) = P(D) \setminus \{\emptyset\}$.

An *effectivity function* (EF) is a function $E : P(N) \to P(P_0(A))$ satisfying: (i) $A \in E(S)$ for all $S \in P_0(N)$, (ii) $E(\emptyset) = \emptyset$, and (iii) $E(N) = P_0(A)$. An effectivity function E is *monotonic* if:

$$[S \in P_0(N),\ S' \supseteq S, \text{ and } B' \supseteq B,\ \ B \in E(S)] \Rightarrow B' \in E(S').$$

An effectivity function E is *superadditive* if:

$$[B_i \in E(S_i),\ i = 1, 2, \text{ and } S_1 \cap S_2 = \emptyset] \Rightarrow B_1 \cap B_2 \in E(S_1 \cup S_2).$$

A *social choice correspondence* (SCC) is a function $H : W^N \to P_0(A)$ where W is the set of *weak* (i.e., complete and transitive) orderings of A.

Assumption: All SCC H considered satisfy: For all $x \in A$ there exists $R^N \in W^N$ such that $H(R^N) = \{x\}$.

Let $H : W^N \to P_0(A)$ be an SCC. A coalition $S \in P_0(N)$ is *effective* for $B \in P_0(A)$ if there exists $Q^S \in W^S$ such that for all $R^{N\setminus S} \in W^{N\setminus S}$, $H(Q^S, R^{N\setminus S}) \subseteq B$. The EF of H, denoted by E^H, is given by $E^H(\emptyset) = \emptyset$ and for $S \in P_0(N)$,

$$E^H(S) = \{B \in P_0(A) | S \text{ is effective for } B\}.$$

By the above assumption, E^H is indeed an EF.

Remark 1 The definition of E^H is valid for any restricted domain SCC, $H : \tilde{W}^N \to P_0(A)$ where $\tilde{W}$ is any nonempty subset of W (that satisfies some mild conditions).

Remark 2 If H and H' are two SCCs such that $H'(R^N) \subseteq H(R^N)$ for all $R^N \in W^N$, then $E^H(S) \subseteq E^{H'}(S)$ for all $S \subseteq N$.

Definition 1 A social choice correspondence H is a *representation* of the effectivity function E if $E^H = E$.

For a finite set D denote by $\Delta(D)$ the set of all probability distributions on D.

A *decision scheme* (DS) is a function $d : W^N \to \Delta(A)$. With a decision scheme d, we associate an SCC which we denote by H_d and define by:

$$H_d(R^N) = \{x \in A | d(x; R^N) > 0\}.$$

A decision scheme d is said to be a *representation* of the effectivity function E if $E^{H_d} = E$. A decision scheme d is said to be *derived* from the social choice correspondence H if $H_d(R^N) = H(R^N)$ for all $R^N \in W^N$.

1.1 The uniform core

The notion of *uniform core* will play an important role in our analysis. Let $E : P(N) \to P(P_0(A))$ be a monotonic and superadditive EF. For any weak preference relation on A, $R \in W$, we denote the strict preference by P, that is, xPy holds for $x, y \in A$ if xRy and not yRx, and the indifference relation by I, that is, xIy holds for $x, y \in A$ if xRy and yRx. Given a profile of preference relations R^N and a coalition $S \subseteq N$, we write $B\ P^S(A \setminus B)$ if xP^iy for all $x \in B$, $y \in A \setminus B$ and $i \in S$.

For $R^N \in W^N$ and an effectivity function E, we define the *uniform core* of E and R^N as follows.

Definition 2 Given an effectivity function E and a profile of preference relations R^N,

- An alternative $x \in A$ is *uniformly dominated* by $B \subseteq A$, $x \notin B$ via the coalition $S \in P_0(N)$, if $B \in E(S)$ and $B\ P^S(A \setminus B)$.
- An alternative $x \in A$ is *not uniformly dominated* at (E, R^N) if there is no pair (S, B) of coalition $S \in P_0(N)$ and a set of states B not containing x that uniformly dominates x via the coalition S.
- The *uniform core* of (E, R^N), denoted by $C_{uf}(E, R^N)$, is the set of all alternatives in A that are not uniformly dominated at (E, R^N).

When the underlying effectivity function E is fixed, we write shortly $C_{uf}(R^N)$ instead of $C_{uf}(E, R^N)$.

This notion is to be compared to the notion of the *core* of an effectivity function defined as follows,

Definition 3 Given an effectivity function E and a profile of preference relations R^N,

- An alternative $x \in A$ is *dominated* by $B \subseteq A$, $x \notin B$ via the coalition $S \in P_0(N)$, if $B \in E(S)$ and $B\ P^S\{x\}$.
- An alternative $x \in A$ is *not dominated* at (E, R^N) if there is no pair (S, B) of a coalition $S \in P_0(N)$ and a set of states B not containing x that dominates x via the coalition S.

– The *core* of (E, R^N), denoted by $C(E, R^N)$, is the set of all alternatives in A that are not dominated at (E, R^N).

It follows from the definitions that the core is a subset of the uniform core. In the following example, based on the Condorcet paradox, the core is empty while the uniform core is not.

Example 2 Let $N = \{1, 2, 3\}$, $A = \{x, y, z\}$ and the effectivity function E given by:

$$E(S) = \begin{cases} P_0(A) & \text{if } |S| > 1 \\ \{A\} & \text{if } |S| = 1 \end{cases}$$

For the profile of preference relations:

$$R^N = \begin{array}{ccc} \underline{1} & \underline{2} & \underline{3} \\ x & z & y \\ y & x & z \\ z & y & x \end{array}$$

At (E, R^N), every alternative is dominated but not uniformly dominated. Hence, $C(E, R^N) = \emptyset$ while $C_{uf}(E, R^N) = A$.

Given a preference profile $R^N = (R^1, \ldots, R^n)$ and a coalition $S \subseteq N$, we denote by $Q^N = (R^S, I^{N\setminus S})$ the preference profile in which $Q^i = R^i$ for $i \in S$ and $R^i = I$ for $i \in N \setminus S$, where I is the total indifference relation on A, that is xIy for all $x, y \in A$.

Remark 3 For any R^N and for any $S \subseteq N$, we have $C_{uf}(R^N) \subseteq C_{uf}(R^S, I^{N\setminus S})$.

Indeed, since uniform domination is defined via strict preference, replacing a strict preference of a player by indifference reduces (weakly) uniform dominance and hence increases (weakly) the uniform core.

As stated in the following theorem, for a monotone and superadditive effectivity function, the uniform core is always nonempty.

Theorem 1 (Abdou and Keiding 1991). *Let E be a monotonic and superadditive EF and let $R^N \in W^N$. Then the uniform core $C_{uf}(E, R^N)$ is nonempty.*

Corollary 1 *For any monotonic and superadditive effectivity function E, the uniform core $C_{uf}(E, \cdot) : W^N \to P_0(A)$ is a social choice correspondence.*

The following result is strongly used in this paper.

Theorem 2 (Keiding and Peleg 2006; Peleg and Peters 2010). *Let E be a monotonic and superadditive EF. Then the social choice correspondence $C_{uf}(E, \cdot)$ is a representation of E, that is $E^{C_{uf}} = E$.*

Corollary 2 *Given a monotonic and superadditive effectivity function E, then any decision scheme d whose support is the uniform core (i.e.,* $\{x \in A | d(x; R^N) > 0\} = C_{uf}(E, R^N)$ *for all* $R^N \in W^N$*), is a representation of E. In particular, any monotonic and superadditive effectivity function has a representation by a decision scheme. For example, the decision scheme denoted by* d_{uf} *and defined by* $d_{uf}(x; R^N) = 1/|C_{uf}(E, R^N)|$ *for* $x \in C_{uf}(E, R^N)$ *and* $d_{uf}(x; R^N) = 0$ *otherwise, which will be called* the uniform representation *of E.*

1.2 Example 1 continued

Using Theorem 2, we know that $C_{uf}(E, \cdot)$ is a representation of E by a social choice correspondence. This can be converted into a representation by a decision scheme (of the same effectivity function) by assigning the uniform distribution on $C_{uf}(E, R^N)$ for each $R^N \in W^N$. For example, let $R^1 = (ww, wb, bw, bb)$ and $R^2 = (bw, wb, ww, bb)$. As can easily be verified $C_{uf}(E, R^N) = \{ww, wb\}$ and hence the uniform decision scheme representing E satisfies $d(ww, R^N) = d(wb, R^N) = 1/2$.

2 Bayes–Nash equilibrium representation

A decision scheme d applied to a situation of a collective choice of a social state induces a game in which each member of the society (player), endowed with a von-Neumann Morgenstern utility function on $\Delta(A)$, chooses a preference relation and the final state is chosen (randomly) according to the decision scheme d. When a player may have incomplete information about the preferences of the other players, this is a game of incomplete information. The question addressed in this paper is:

Given a monotone and superadditive effectivity function E, can it be represented by a decision scheme so that the induced game of incomplete information has a Bayes Nash equilibrium in pure strategies?

We provide an affirmative answer to this question. For the sake of the presentation, we will first state and prove the result for the situation of complete information and then state and prove the more general result for the incomplete information situation.

2.1 The complete information case

Given a society $N = \{1, 2, \ldots, n\}$, a set of social states $A = \{a_1, a_2, \ldots, a_m\}$, and effectivity function E, a utility function of player i is a von-Neumann Morgenstern utility function on $\Delta(A)$, induced by $u^i : A \to \mathbb{R}$. For any decision scheme $d : W^N \to \Delta(A)$ consider the strategic form game $\Gamma_d = (N; W, \ldots, W; u^1, \ldots, u^n; d)$. This is the strategic form game in which the set of players is N and the set of actions (pure strategies) of each player $i \in N$ is W, the set of weak orderings of the social states A. Given a vector $R^N \in W^N$ of pure actions chosen by the players, any social state $x \in A$ is chosen (by the decision scheme) with probability $d(x; R^N)$. The payoff of each player i is his (von-Neumann Morgenstern) utility $u^i(x)$. Our objective is to find a decision scheme d representing the effectivity function E such that the game Γ_d has a NE in pure strategies. We illustrate this in the following example.

Example 3 (*Neutral effectivity functions*) A *veto function* is a function $v : P(N) \to \{-1, 0, \ldots, m-1\}$ such that $v(\emptyset) = -1$, $v(S) \geq 0$ if $S \neq \emptyset$ and $v(N) = m - 1$. The interpretation is that a nonempty set of players S, can veto any set of at most $v(S)$ alternatives. A veto function v defines a neutral EF, $E_v : P(N) \to P(P_0(A))$ by $E_v(\emptyset) = \emptyset$ and $E_v(S) = \{B | v(S) \geq m - |B|\}$ for $S \neq \emptyset$. That is, $B \in E_v(S)$ if the coalition S can veto all the alternatives in $A \setminus B$. This effectivity function is neutral (with respect to the alternatives) as $E_v(S)$ does not depend on the names of the alternatives. We remark that E_v is monotonic (superadditive) if and only if v is monotonic (superadditive). We assume complete information. Thus, the specification of the model is completed by an n-tuple of utility functions for the players: $u^i : A \to \mathbb{R}, \quad i = 1, \ldots, n$.

Let $E : P(N) \to P(P_0(A))$ be a monotonic , superadditive, and neutral EF. Let $v : P(N) \to \{-1, 0, \ldots, m-1\}$ be the veto function of E and let $R^N \in W^N$. *Sincere vetoing* with respect to v and R^N (in the natural ordering of the players) is as follows: Player 1 vetoes $v(1)$ of his worst alternatives; next, player 2 vetoes $v(2)$ of his worst alternatives in the remaining set of alternatives and so forth. By superadditivity, $v(1) + \cdots + v(n) \leq v(N) = m - 1$ hence there is always a nonempty set of remaining alternatives. Clearly, this could be done with any other ordering of the players.

By Peleg and Peters (2010, Theorem 6.4.4), for any preference ordering profile R^N, sincere vetoing (with respect to any ordering of the players) is a Nash equilibrium for the uniform decision scheme d_{uf} (see Corollary 2). The following is an explicit example.

Let $N = \{1, 2, 3\}$ and $A = \{a, b, c, d, e\}$. Let the veto function v be specified by $v(i) = 1$, $v(i, j) = 2$ for all i and j and $v(N) = 4$. Let the preferences be defined through the utility functions:

$$u^1(a) > u^1(b) > u^1(c) > u^1(d) > u^1(e),$$
$$u^2(e) > u^2(d) > u^2(c) > u^2(b) > u^2(a), \text{ and}$$
$$u^3(a) > u^3(e) > u^3(b) > u^3(d) > u^3(c).$$

Vetoing sincerely amounts to presenting the following dichotomous preferences (see Sect. 2.5 for the notation):

$$R^1 = \frac{a\,b\,c\,d}{e} \quad R^2 = \frac{b\,c\,d\,e}{a} \quad R^3 = \frac{d\,b\,e\,a}{c}$$

The profile of preference orderings $R^N = (R^1, R^2, R^3)$ is indeed the desired NE (note that $C_{uf}(R^N) = \{b, d\}$).

The general result of this kind is given by the following theorem.

Theorem 3 *Given a monotonic and superadditive effectivity function E, and vNM utility functions* $(u^1, \ldots, u^n)$, *then there is a decision scheme d such that,*

- *The decision scheme d is a representation of the effectivity function E.*

– *The game $\Gamma_d = (N; W, \dots, W; u^1, \dots, u^n; d)$ has a Nash equilibrium in pure strategies.*

Proof Let $d_{uf} : W^N \to \Delta(A)$ be uniform core representation of E defined in Corollary 2. Let $q = (q(s))_{s \in S}$ be correlated equilibrium of the game $\Gamma_{d_{uf}} = (N; W, \dots, W; u^1, \dots, u^n; d_{uf})$ (where $S = W \times \dots \times W$ denotes the set of vectors of pure strategies in this game)). Then,

$$\sum_{s \in S} q(s) \sum_{x \in A} u^i(x) d_{uf}(x; s) \geq \sum_{s \in S} q(s) \sum_{x \in A} u^i(x) d_{uf}(x; (s^{-i}, R^i)), \tag{1}$$

holds for all $i \in N$ and $R^i \in W$.

Define now a new decision scheme d by:

1. $d(x; I, \dots, I) = \sum_{s \in S} q(s) d_{uf}(x; s), \quad \forall x \in A.$
2. $d(x; \overbrace{I, \dots, I}^{i-1}, R^i, \overbrace{I, \dots, I}^{n-i}) = \sum_{s \in S} q(s) d_{uf}(x; (s^{-i}, R^i)), \quad \forall x \in A, \ \forall i \in N, \ \forall R^i \in W.$
3. Otherwise, for any other profile of preferences $R^N \in W^N$ and any $x \in A$ define $d(x; R^N) = d_{uf}(x; R^N)$.

By its definition, the support of d is contained in the uniform core, that is, the induced social choice correspondence H_d satisfies $H_d(R^N) \subseteq C_{uf}(R^N)$ for all $R^N \subseteq W^N$. By Remark 2, $E^{H_d}(S) \supseteq E^{C_{uf}}(S)$ for all $S \subseteq N$. By Theorem 2, $E^{C_{uf}} = E$ and hence $E^{H_d}(S) \supseteq E(S)$ for all $S \subseteq N$.

In order to prove that d is a representation of E, all we have to show is that part 2 in the definition of d does not give extra power (w.r.t. d) to $N \setminus \{i\}$ for any player i. Suppose $N \setminus \{i\}$ is not effective for B (according to E), then by part 3 of the definition of d, $N \setminus \{i\}$ is not effective for B via any strategy vector different from I^{-i}. It remains to see that $N \setminus \{i\}$ cannot guarantee an outcome in B by the strategy vector I^{-i}. Indeed, choose a strategy vector s such that $q(s) > 0$, then choose $x \notin B$ and $R^i \in W$ such that $d_{uf}(x; (s^{-i}, R^i)) > 0$ (such x and R^i exist since $N \setminus \{i\}$ is not effective for B according to E). Then, part 2 in the definition of d implies that $d(x; (I^{-i}, R^i)) > 0$ and thus, $N \setminus \{i\}$ is not effective for B w.r.t. d via I^{-i}.

Inequalities (1) imply that the pure strategy vector $(I, \dots, I)$ is a Nash equilibrium in the game Γ_d. Indeed, for any deviation $R^i \in W$ of player i,

$$\begin{aligned}
\sum_{x \in A} u^i(x) d(x; I, \dots, I) &= \sum_{x \in A} u^i(x) \sum_{s \in S} q(s) d_{uf}(x; s) \\
&= \sum_{s \in S} q(s) \sum_{x \in A} u^i(x) d_{uf}(x; s) \\
&\geq \sum_{s \in S} q(s) \sum_{x \in A} u^i(x) d_{uf}(x; s^{-i}, R^i) \\
&= \sum_{x \in A} u^i(x) \sum_{s \in S} q(s) d_{uf}(x; s^{-i}, R^i)
\end{aligned}$$

$$= \sum_{x \in A} u^i(x) d(x; \overbrace{I, \ldots, I}^{i-1}, R^i, \overbrace{I, \ldots, I}^{n-i})$$

□

2.2 The incomplete information case: main result

An *information structure* (IS) is a $2n$-tuple $\mathcal{I} = (T^1, \ldots, T^n; p^1, \ldots, p^n)$ where for each $i \in N$, T^i is the (finite) set of *types* of player i, and p^i is a probability distribution on $T = \times_{i \in N} T^i$ such that $p^i(t^i = t_0^i) > 0$ for all $t_0^i \in T^i$. This is the *prior distribution* of player i on the set of types T, which induces the conditional probability distribution $p^i(t^{-i}|t^i)$ on $T^{-i} = \times_{j \neq i} T^j$ (the beliefs of player i of type t^i on the types of the other players). In a Harsanyi-consistent information structure there is a *common prior* namely, $p^i = p$, for all $i \in N$.

We now modify the notion of decision scheme so as to adapt it to the context of incomplete information.

Definition 4 1. A generalized decision scheme (GDS) is a function $d : W^N \times T \to \Delta(A)$.

2. A strategy of player i (with respect to a GDS) is a pair (s^i, π^i) where $s^i : T^i \to W$ (denote by S^i the set of all such mappings, let $S = S^1 \times \cdots \times S^n$) and $\pi^i : T^i \to T^i$. Equivalently, a strategy of player i is a mapping $\tilde{s}^i : T^i \to W \times T^i$. Denote by $\tilde{S}^i$ the set of pure strategies of player i and by $\tilde{S} = \tilde{S}^1 \times \cdots \times \tilde{S}^n$ the set of vectors of pure strategies. A vector $\tilde{s} \in \tilde{S}$ will also be written as $\tilde{s} = (s, \pi)$ where $s = (s^1, \ldots, s^n) \in S$ and $\pi = (\pi^1, \ldots, \pi^n)$.

The idea behind this definition is that in a situation of incomplete information, each player is asked to input to the (generalized) decision scheme, both his preferences and his type. As a result, the (Bayes Nash) equilibrium of the induced game will exhibit the 'spirit' of the *revelation principle*, as each player will input his true type.

Any generalized decision scheme (GDS) induces an effectivity functions in a similar way that a DS does. Let $d : W^N \times T \to \Delta(A)$ be a GDS. The associated (generalized) SCC, $H : W^N \times T \to P_0(A)$ is defined by $H(R^N, t) = \{x \in A | d(x; R^N, t) > 0\}$.

A coalition S is effective for a nonempty subset B of A (w.r.t. H) if there exist $R^S \in W^S$ and $t^S \in T^S$ such that $H(R^S, Q^{N \setminus S}, t^S, r^{N \setminus S}) \subseteq B$ for all $Q^{N \setminus S} \in W^{N \setminus S}$ and $r^{N \setminus S} \in T^{N \setminus S}$. The effectivity function of H is defined by $E^H(S) = \{B | S \text{ is effective for } B\}$. The effectivity function of d is defined to be that of H. The generalized decision scheme d is a representation of a given (monotonic and superadditive) effectivity function E if the effectivity function of d equals E.

Let $\mathcal{I} = (T^1, \ldots, T^n; p^1, \ldots, p^n)$ be an IS and let $(u^i)_{i \in N}$ where $u^i : A \times T \to \mathbb{R}$, be the vector of utility functions of the players. Then, a generalized decision scheme d defines a Bayesian game of incomplete information:

$$\Gamma_{Id} = (N; W, \ldots, W; \mathcal{I}; u^1, \ldots, u^n; d).$$

This is the strategic form game in which:

- The set of actions of player $i \in N$ of any possible type t^i is $W \times T^i$. The set of pure strategies of player i is $\tilde{S}^i$.
- The payoff to type t^i when the players play the pure strategy vector $\tilde{s} = (\tilde{s}^1, \dots, \tilde{s}^n) \in \tilde{S}$ is $U^i(\tilde{s}|t^i)$ given by:

$$U_d^i(\tilde{s}|t^i) = \sum_{t^{-i} \in T^{-i}} p^i(t^{-i}|t^i) \sum_{x \in A} u^i(x;t) d(x; \tilde{s}^1(t^1), \dots, \tilde{s}^n(t^n)). \quad (2)$$

When $d(\cdot; R^N, t)$ does not depend on t, the expected utility $U_d^i(\tilde{s}|t^i)$ also does not depend on t and we write:

$$U_d^i(s|t^i) = \sum_{t^{-i} \in T^{-i}} p^i(t^{-i}|t^i) \sum_{x \in A} u^i(x;t) d(x; s^1(t^1), \dots, s^n(t^n)). \quad (3)$$

As expected, the dependence of $U_d^i(s|t^i)$ on s^i is only via $s^i(t^i)$ and it does not depend on $s^i(\hat{t}^i)$ for $\hat{t}^i \neq t^i$.
The difference between the incomplete information game Γ_{Id} and the complete information game Γ_d defined in page 5 is the fact that d is now a generalized decision scheme and the addition of the information structure $\mathscr{I}$ which imply that the action set of player i is not W but rather a mapping from his type set T^i to $W \times T^i$.

An n-tuple of strategies $\tilde{s}$ is a *Bayesian Nash equilibrium* (BNE) if for all $i \in N$, all $t^i \in T^i$ and all $(R^i, \hat{t}^i) \in W \times T^i$,

$$\sum_{t^{-i} \in T^{-i}} p^i(t^{-i}|t^i) \sum_{x \in A} u^i(x;t) d(x; \tilde{s}(t)) \geq \sum_{t^{-i} \in T^{-i}} p^i(t^{-i}|t^i) \sum_{x \in A} u^i(x;t) \times d((x; \tilde{s}^{-i}(t^{-i}), (R^i, \hat{t}^i))), \quad (4)$$

where $\tilde{s}(t)$ is the vector $(\tilde{s}^i(t^i))_{i \in N}$ and $\tilde{s}^{-i}(t^{-i})$ is the vector $(\tilde{s}^j(t^j))_{j \neq i}$.

Theorem 4 *Let $E : P(N) \to P(P_0(A))$ be a monotonic and superadditive EF. Let $\mathscr{I} = (T^1, \dots, T^n; p^1, \dots, p^n)$ be an IS, and let $(u^1, \dots, u^n)$, $u^i : A \times T \to \mathbb{R}$, be a vector of vNM utilities for the players. Then E has a representation by a generalized decision scheme $d : W^N \times T \to \Delta(A)$ such that the game $\Gamma_{Id} = (N; W, \dots, W; \mathscr{I}; (u^i)_{i \in N}; d)$ has a BNE in pure strategies in which all players report their true types.*

Proof Define the generalized decision scheme $d_1 : W^N \times T \to \Delta(A)$ by

$$d_1(R^N, t) = d_{uf}(R^N), \quad \forall (R^N, t) \in W^N \times T.$$

As $d_1(R^N, t)$ depends only on R^N, by slight abuse of notation, we shall also write $d_1(R^N)$ instead of $d_1(R^N, t)$. Consider now the ex-ante game:

$$G_{d_1} = (N; S^1, \ldots, S^n; h^1, \ldots, h^n; d_1) \tag{5}$$

in which the payoff functions are:

$$h^i(s^1, \ldots, s^n) = \sum_{t \in T} p^i(t) \sum_{x \in A} u^i(x, t) d_1(x; s(t)), \tag{6}$$

which can be written as (by (3) as d_1 does not depend on π):

$$\begin{aligned} h^i(s^1, \ldots, s^n) &= \sum_{t^i \in T^i} p^i(t^i) \sum_{t^{-i} \in T^{-i}} p^i(t^{-i}|t^i) \sum_{x \in A} u^i(x, t) d_1(x; s(t)) \\ &= \sum_{t^i \in T^i} p^i(t^i) U_d^i(s|t^i). \end{aligned} \tag{7}$$

Note that in this game, the strategy sets are S^i rather than $\tilde{S}^i$ since $d_1(R^N, t)$ does not depend on t. This is a finite game with complete information, so it has a correlated equilibrium (CE). Let $(q(s))_{s \in S}$ be a CE of the game G_{d_1}, then the equilibrium condition is:

$$\sum_{s \in S} q(s) h^i(s) \geq \sum_{s \in S} q(s) h^i(s^{-i}, \delta^i(s^i)), \tag{8}$$

which holds for all $i \in N$ and for all $\delta^i : S^i \to S^i$. Substituting h^i from (7) we have:

$$\sum_{s \in S} q(s) \sum_{\hat{t}^i \in T^i} p^i(\hat{t}^i) U_{d_1}^i(s|\hat{t}^i) \geq \sum_{s \in S} q(s) \sum_{\hat{t}^i \in T^i} p^i(\hat{t}^i) U_{d_1}^i(s^{-i}, \delta^i(s^i)|\hat{t}^i),$$

which we rewrite as:

$$\sum_{\hat{t}^i \in T^i} p^i(\hat{t}^i) \sum_{s \in S} q(s) U_{d_1}^i(s|\hat{t}^i) \geq \sum_{\hat{t}^i \in T^i} p^i(\hat{t}^i) \sum_{s \in S} q(s) U_{d_1}^i(s^{-i}, \delta^i(s^i)|\hat{t}^i). \tag{9}$$

For $t^i \in T^i$ let $\delta^i : S^i \to S^i$ be defined as follows:

- $\delta^i(s^i)(t^i) = R^i \in W, \ \forall s^i \in S^i$.
- $\delta^i(s^i)(\hat{t}^i) = s^i(\hat{t}^i)$, if $\hat{t}^i \neq t^i$, for all $s^i \in S^i$.

Inserting this δ^i in (9) all terms with $\hat{t}^i \neq t^i$ will be the same on both sides of the inequality and will cancel, dividing the remaining term by $p^i(t^i)$ (which is positive) we obtain that:

$$\sum_{s \in S} q(s) U_{d_1}^i(s|t^i) \geq \sum_{s \in S} q(s) U_{d_1}^i(s^{-i}, R^i|t^i) \tag{10}$$

holds for all $i \in N$, $t^i \in T^i$ and $R^i \in W$.

For $t^i \in T^i$ and $\tilde{t}^i \in T^i$ let $\tilde{\delta}^i : S^i \to S^i$ be defined as follows:

- $\tilde{\delta}^i(s^i)(t^i) = s^i(\tilde{t}^i), \quad \forall s^i \in S^i$.
- $\tilde{\delta}^i(s^i)(\hat{t}^i) = s^i(\hat{t}^i)$, if $\hat{t}^i \neq t^i$, for all $s^i \in S^i$.

Inserting this $\tilde{\delta}^i$ in (9), all terms with $\hat{t}^i \neq t^i$ will be the same on both sides of the inequality and will cancel, dividing the remaining term by $p^i(t^i)$ (which is positive) we obtain that:

$$\sum_{s \in S} q(s) U^i_{d_1}(s|t^i) \geq \sum_{s \in S} q(s) U^i_{d_1}(s^{-i}, s^i(\tilde{t}^i)|t^i) \tag{11}$$

holds for all $i \in N$ and for all t^i and $\tilde{t}^i$ in T^i.

Define now a generalized decision scheme d by:

$$d(x; I^N, t) = \sum_{s \in S} q(s) d_1(x; s(t)), \ \forall x \in A, \ \forall t \in T. \tag{12}$$

$$d(x; (I^{-i}, R^i), t) = \sum_{s \in S} q(s) d_1(x; s^{-i}(t^{-i}), R^i), \tag{13}$$

$$\forall i \in N, \ R^i \in W, \ R^i \neq I, \ t \in T, \ x \in A.$$

$$d(x; R^N, t) = d_{uf}(x; R^N) \quad \text{otherwise.} \tag{14}$$

We first claim that d is a representation of the effectivity function E. The idea of the proof is the same as in the proof of Theorem 3: As argued there (see page 6), by Remark 2 and Theorem 2, the effectivity function of d is at least as rich as E. We have to show that (13) does not give extra power, w.r.t. d, to $N \setminus \{i\}$ for every $i \in N$. Suppose $N \setminus \{i\}$ is not effective for B (according to E), then by part (14) of the definition of d, $N \setminus \{i\}$ is not effective for B via any strategy vector different form I^{-i}. It remains to see that $N \setminus \{i\}$ cannot guarantee an outcome in B by the strategy vector I^{-i}. Indeed, choose a strategy vector s such that $q(s) > 0$, then for every $t \in T$ choose $x \notin B$ and $R^i \in W$ such that $d_{uf}(x; (s^{-i}(t^{-i}), R^i)) > 0$ (such x and R^i exist since $N \setminus \{i\}$ is not effective for B according to E). Then, part (13) in the definition of d implies that $d(x; (I^{-i}, R^i, t)) > 0$ and thus, at any $t \in T$, $N \setminus \{i\}$ is not effective for B w.r.t. d via I^{-i}.

We next claim that the pure strategy vector $\tilde{s}$ in which $\tilde{s}^i(t^i) = (I, t^i)$, for all $i \in N$ and for all $t^i \in T^i$, is a BNE of the game $\Gamma_{Id} = (N; W, \ldots, W; \mathscr{I}; (u^i)_{i \in N}; d)$ that is, inequalities (4) are satisfied for any $t^i \in T^i$, $\tilde{s}^i(t^i) = (I, t^i)$ and any deviation to $(R^i, \hat{t}^i)$. To do this, we shall treat each of the three possible deviations:

(i) Deviation from (I, t^i) to (R^i, t^i) with $R^i \neq I$.
(ii) Deviation from (I, t^i) to $(I, \tilde{t}^i)$ with $\tilde{t}^i \neq t^i$.
(iii) Deviation from (I, t^i) to $(R^i, \tilde{t}^i)$ with $R^i \neq I$ and $\tilde{t}^i \neq t^i$.

Case (i). Substituting $\tilde{s} = (I^N, t)$ in the left-hand side of (4) and d from(12) we have:

$$\begin{aligned}
&\sum_{t^{-i}\in T^{-i}} p^i(t^{-i}|t^i)\sum_{x\in A} u^i(x;t)d(x;I^N,t)\\
&\qquad = \sum_{t^{-i}\in T^{-i}} p^i(t^{-i}|t^i)\sum_{x\in A} u^i(x;t)\sum_{s\in S} q(s)d_1(x;s(t))\\
&\qquad = \sum_{s\in S} q(s)\sum_{t^{-i}\in T^{-i}} p^i(t^{-i}|t^i)\sum_{x\in A} u^i(x;t)d_1(x;s(t))\\
\text{by (3)}\quad &\qquad = \sum_{s\in S} q(s)U^i_{d_1}(s|t^i)\\
\text{by (10)}\quad &\qquad \geq \sum_{s\in S} q(s)U^i_{d_1}(s^{-i},R^i|t^i)\\
\text{by (3)}\quad &\qquad = \sum_{s\in S} q(s)\sum_{t^{-i}\in T^{-i}} p^i(t^{-i}|t^i)\sum_{x\in A} u^i(x;t)d_1(x;s^{-i}(t^{-i}),R^i)\\
\text{by (13)}\quad &\qquad = \sum_{t^{-i}\in T^{-i}} p^i(t^{-i}|t^i)\sum_{x\in A} u^i(x;t)d(x;(I^{-i},R^i),t).
\end{aligned}$$

Case (ii). Substituting $\tilde{s} = (I^N, t)$ in the left-hand side of (4) and d from(12) we have:

$$\begin{aligned}
&\sum_{t^{-i}\in T^{-i}} p^i(t^{-i}|t^i)\sum_{x\in A} u^i(x;t)d(x;I^N,t)\\
&\qquad = \sum_{t^{-i}\in T^{-i}} p^i(t^{-i}|t^i)\sum_{x\in A} u^i(x;t)\sum_{s\in S} q(s)d_1(x;s(t))\\
&\qquad = \sum_{s\in S} q(s)\sum_{t^{-i}\in T^{-i}} p^i(t^{-i}|t^i)\sum_{x\in A} u^i(x;t)d_1(x;s(t))\\
\text{by (3)}\quad &\qquad = \sum_{s\in S} q(s)U^i_{d_1}(s|t^i)\\
\text{by (11)}\quad &\qquad \geq \sum_{s\in S} q(s)U^i_{d_1}(s^{-i},s^i(\tilde{t}^i)|t^i)\\
\text{by (3)}\quad &\qquad = \sum_{s\in S} q(s)\sum_{t^{-i}\in T^{-i}} p^i(t^{-i}|t^i)\sum_{x\in A} u^i(x;t)d_1(x;s^{-i}(t^{-i}),s^i(\tilde{t}^i))\\
\text{by (12)}\quad &\qquad = \sum_{t^{-i}\in T^{-i}} p^i(t^{-i}|t^i)\sum_{x\in A} u^i(x;t)d(x;I^N,(t^{-i},\tilde{t}^i)).
\end{aligned}$$

Case (iii). This case follows from case (i) since by (13), for $R^i \neq I$:

$$\begin{aligned}
&d(x;(I^{-i},R^i),(t^{-i},\tilde{t}^i)) = d(x;(I^{-i},R^i),t)\\
&\quad = \sum_{s\in S} q(s)d_1(x;s^{-i}(t^{-i}),R^i),\ \forall i\in N,\ R^i\in W,\ t\in T,\ x\in A.
\end{aligned}$$

□

2.3 Monotonic decision schemes

Monotonicity is an intuitive and desirable property of a decision scheme representing a certain EF. It roughly says that if the position of an alternative is improved in the preferences of the members the society, its probability of being chosen by the decision scheme should increase. This requirement may be too strong when the alternative is ranked very low in the members' preferences so that its probability of being chosen is zero and it is likely to remain zero even if one member improves its ranking in his preference. To express this formally, we first introduce the following term and notation:

Let $R^N \in W^N$ and let $x \in A$ be any alternative. A preference profile $\tilde{R}^N \in W^N$ is said to be *obtained from R^N by an improvement of the position of x* if:

1. $aR^i b \Leftrightarrow a\tilde{R}^i b$ for all $a, b \in A \setminus \{x\}$ and all $i \in N$.
2. $xR^i a \Rightarrow x\tilde{R}^i a$ for all $a \neq x$ and all $i \in N$.
3. There exists $j \in N$ and $y \in A$ such that $yR^j x$ and $x\tilde{P}^j y$ or $yP^j x$ and $x\tilde{R}^j y$.

We denote this by $R^N \Uparrow x\ \tilde{R}^N$.

Example 4 Let $N = \{1, 2\}$, $A = \{a, b, c\}$, $x = c$.

If $R^N = \begin{array}{cc} \underline{R^1} & \underline{R^2} \\ ab & a \\ c & b \\ & c \end{array}$ and $\tilde{R}^N = \begin{array}{cc} \underline{\tilde{R}^1} & \underline{\tilde{R}^2} \\ abc & a \\ & b \\ & c \end{array}$ then $R^N \Uparrow x\ \tilde{R}^N$.

Using this notation and incorporating the above-mentioned weakening of the monotonicity property, we are led to the following definition:

Definition 5 A decision scheme, $d : W^N \to \Delta(A)$, is *weakly monotonic* if for all $x \in A$ and $R^N, \tilde{R}^N \in W^N$ such that $R^N \Uparrow x\ \tilde{R}^N$,

1. $d(x, \tilde{R}^N) \geq d(x, R^N)$ and
2. $d(y, R^N) = 0 \Rightarrow d(y, \tilde{R}^N) = 0$ for all $y \neq x$.

Definition 6 A decision scheme, $d : W^N \to \Delta(A)$, is (strictly) *monotonic* if it is weakly monotonic and if $0 < d(x, R^N) < 1$ and $R^N \uparrow x\ \tilde{R}^N$ then $d(x, \tilde{R}^N) > d(x, R^N)$.

Note that the decision scheme d_{uf} which is the uniform probability distribution over the uniform core is weakly monotonic but it is not monotonic, since improving the position of an alternative may leave the uniform core unchanged and hence the probabilities assigned by d_{uf} are unchanged. As we shall apply these notions of monotonicity also for generalized decision schemes, we define:

Definition 7 A generalized decision scheme: $d : W^N \times T \to \Delta(A)$ is monotonic (resp. weakly monotonic) if for any $t \in T$ the decision scheme $d_t : W^N \to \Delta(A)$ defined by $d_t(R^N) = d(R^N; t)$ for all $R^N \in W^N$ is monotonic (resp. weakly monotonic).

For our results so far, we heavily used the fact that for every monotonic and superadditive effectivity function E, the decision scheme d_{uf} is a representation. In view of the fact that this decision scheme is not monotonic, a natural question is whether we can always find another representation by a monotonic decision scheme ? We provide a positive answer in the following

Theorem 5 *Any monotonic and superadditive effectivity function E has a representation by a monotonic decision scheme.*

Proof For $R^N \in W^N$ and $x \in A$, denote by $B_d(x, R^N)$ the Borda count (for weak orderings)[1] of x with respect to R^N. Define a decision scheme $d^*_{uf} : W^N \to \Delta(A)$ by:

1. $d^*_{uf}(x, R^N) = 0$, if $d_{uf}(x, R^N) = 0$.
2. $d^*_{uf}(x, R^N) = \frac{B_d(x,R^N)}{\sum_{\{y|d_{uf}(y,R^N)>0\}} B_d(y,R^N)}$, if $d_{uf}(x, R^N) > 0$.

Thus, $d^*_{uf}(\cdot, R^N)$ has the same support as $d_{uf}(\cdot, R^N)$, namely the uniform core of R^N but rather than the uniform probability distribution in $d_{uf}(\cdot, R^N)$, the probabilities assigned by the decision scheme $d^*_{uf}(\cdot, R^N)$ take into account the relative ranking of the uniform core elements. So, d^*_{uf} is a representation of E since it has the same support as d_{uf}. To see that it is monotonic, let $x \in A$, $R^N, \tilde{R}^N \in W^N$, $0 < d^*_{uf}(x, R^N) < 1$ and $R^N \uparrow_x \tilde{R}^N$. Then $B_d(x, \tilde{R}^N) > B_d(x, R^N)$ and

$$\gamma_1 := \sum_{\{y|y \neq x \ \& \ d^*_{uf}(y,\tilde{R}^N)>0\}} B_d(y, \tilde{R}^N) \leq \sum_{\{y|y \neq x \ \& \ d^*_{uf}(y,R^N)>0\}} B_d(y, R^N) := \gamma_2.$$

Thus,

$$\begin{aligned} d^*_{uf}(x, \tilde{R}^N) &= \frac{B_d(x, \tilde{R}^N)}{B_d(x, \tilde{R}^N) + \gamma_1} \geq \frac{B_d(x, \tilde{R}^N)}{B_d(x, \tilde{R}^N) + \gamma_2} \\ &> \frac{B_d(x, R^N)}{B_d(x, R^N) + \gamma_2} = d^*_{uf}(x, R^N) \end{aligned}$$

□

We observe that in the proof of Theorems 3 and 4 hold if we replace d_{uf} by d^*_{uf}.

We address now a natural question regarding the proof of our main theorem: Was it indeed necessary to use a generalized decision scheme in order to obtain our representation result rather than a decision scheme that was sufficient for the representation in the complete information case? More precisely, can Theorem 4 be proved with DS

[1] The Borda count is originally defined for strict orderings, however, it can be extended to weak preference orderings as follows: Given $R \in W$ and $x \in A$, the Borda count of x is the average of the Borda counts of the elements of its equivalence class (in R) in a strict preference ordering Q on A that preserves the strict preferences in R, that is if x is strictly preferred to y according to R then $x Q y$ (it is easily verified that this is well defined that is, it is the same for any such Q).

rather than with GDS? In view of Theorem 5 and the restriction of our attention to monotone decision schemes the question is: can Theorem 4 be proved with a *monotone* DS rather than with (a monotone) GDS? The answer to this question is NO which we prove by the following counterexample of an effectivity function (monotone and superadditive) and vNM utility functions for which there is no monotonic representation that generates a game with pure strategy BNE.

Example 5 Let $N = \{1, 2\}$ be the set of players, $A = \{a, b\}$ be the set of states and the effectivity function E is the two-person unanimity EF that is, $E(1) = E(2) = \{A\}$ and $E(N) = P_0(A)$. Assume the type sets are: $T^1 = \{t_1^1, t_2^1, t_3^1, t_4^1, t_5^1, t_6^1\}$ and $T^2 = \{t_1^2, t_2^2, t_3^2\}$ and there is a common prior p which is the uniform prior on $T = T^1 \times T^2$. The utility functions $u^i : A \times T \to \mathbb{R}$ will be specified later.

Claim. There exist utility functions u^1 and u^2 such that for any representation of E by a monotonic decision scheme d, the game

$$\Gamma_{Id} = (N; W, W; T^1, T^2, p; u^1, u^2; d)$$

has no BNE in pure strategies.

Proof (i) Choose $u^i(b, t_k^1, t_\ell^2) = 0$ for all i, k and ℓ.
(ii) Choose u^2 so that $u^2(a, t_k^1, t_\ell^2) > 0\ \forall k, \ell$.
(iii) Choose $u^1(a, t_{3+r}^1, t_\ell^2) = -u^1(a, t_r^1, t_\ell^2)$ for $r = 1, 2, 3$ and $\ell = 1, 2, 3$.
(iv) Choose $u^1(a, t_k^1, t_\ell^2)$ so that the following determinant will be nonzero:

$$\begin{vmatrix} u^1(a, t_1^1, t_1^2) & u^1(a, t_1^1, t_2^2) & u^1(a, t_1^1, t_3^2) \\ u^1(a, t_2^1, t_1^2) & u^1(a, t_2^1, t_2^2) & u^1(a, t_2^1, t_3^2) \\ u^1(a, t_3^1, t_1^2) & u^1(a, t_3^1, t_2^2) & u^1(a, t_3^1, t_3^2) \end{vmatrix} \neq 0 \qquad (15)$$

In the sequel, we shall refer to these properties of u^i as (i), (ii) etc.

Assume now, contrary to our claim that there exists a representation of E by a monotonic decision scheme d such that the game

$$\Gamma_{Id} = (N; W, W; T^1, T^2, p; u^1, u^2; d)$$

has a pure strategy BNE, say $s = (s^1, s^2)$. This implies (using the uniform common prior) the following inequalities:

$$\Sigma_{\ell=1}^3 u^1(a, t_k^1, t_\ell^2) d(a, s^1(t_k^1), s^2(t_\ell^2)) \geq \Sigma_{\ell=1}^3 u^1(a, t_k^1, t_\ell^2) d(a, R, s^2(t_\ell^2)), \quad (16)$$

for $k = 1, \ldots, 6$ and $R \in W \setminus \{s^1(t_k^1)\}$.

Using property (iii), this implies:

$$\Sigma_{\ell=1}^3 u^1(a, t_k^1, t_\ell^2) d(a, s^1(t_k^1), s^2(t_\ell^2)) = \Sigma_{\ell=1}^3 u^1(a, t_k^1, t_\ell^2) d(a, R, s^2(t_\ell^2)), \quad (17)$$

for $k = 1, \dots, 6$ and $R \in W$, and hence,

$$\Sigma_{\ell=1}^{3} u^1(a, t_k^1, t_\ell^2)[d(a, R, s^2(t_\ell^2)) - d(a, Q, s^2(t_\ell^2))] = 0, \qquad (18)$$

for all R and Q in W. By (15) this implies:

$$d(a, R, s^2(t_\ell^2)) = d(a, Q, s^2(t_\ell^2)), \ \forall R, Q \in W, \ \ell = 1, 2, 3. \qquad (19)$$

The equilibrium conditions for player 2 of type t_1^2 are

$$\Sigma_{k=1}^{6} u^2(a, t_k^1, t_1^2) d(a, s^1(t_k^1), s^2(t_1^2)) \geq \Sigma_{k=1}^{6} u^2(a, t_k^1, t_1^2) d(a, s^1(t_k^1), R), \quad (20)$$

for all $R \in W \setminus \{s^2(t_1^2)\}$.

We finish the proof by showing that there is no value of $s^2(t_1^2)$ that satisfies (20). Indeed, if $s^2(t_1^2) = \begin{smallmatrix} a \\ b \end{smallmatrix}$ then by (19) we have $d\left(a, R, \begin{smallmatrix} a \\ b \end{smallmatrix}\right) = \alpha, \forall R \in W$. Since d is a representation of the effectivity function E and $\{1, 2\}$ are effective for $\{a\}$ we must have $\alpha = 1$ since $d\left(a, \begin{smallmatrix} a \\ b \end{smallmatrix}, \begin{smallmatrix} a \\ b \end{smallmatrix}\right) = 1$ and hence $d\left(a, R, \begin{smallmatrix} a \\ b \end{smallmatrix}\right) = 1, \forall R \in W$ which is impossible since $\{a\} \notin E(2)$. Similarly $s^2(t_1^2) = \begin{smallmatrix} b \\ a \end{smallmatrix}$ is impossible. If $s^2(t_1^2) = ab$, we have by monotonicity that $d(a, s^1(t_k^1), s^2(t_1^2)) \leq d(a, s^1(t_k^1), \begin{smallmatrix} a \\ b \end{smallmatrix})$ holds for $k = 1, \dots, 6$ and the inequality is strict when the left-hand side is strictly between 0 and 1. Indeed this must be the case for all k since, using by (19), $d(a, s^1(t_k^1), s^2(t_1^2) = 0$ implies that $\{b\} \in E(2)$ while $d(a, s^1(t_k^1), s^2(t_1^2) = 1$ implies that $\{a\} \in E(2)$, both contradict the representation. Therefore, $d(a, s^1(t_k^1), s^2(t_1^2)) < d\left(a, s^1(t_k^1), \begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ holds for all k and hence (20) is violated, completing the proof. □

2.4 Bayesian incentive compatibility

In this section, we reconsider the incomplete information game with information structure $\mathscr{I} = (T^1, \dots, T^n; p^1, \dots, p^n)$,

$$\Gamma_{Id} = (N; W, \dots, W; \mathscr{I}; u^1, \dots, u^n; d)$$

introduced in a previous section. We assume, as in d'Aspremont and Peleg (1988), that the types of the players include (explicitly) information on their ordinal preferences on the set A of alternatives. More precisely, we assume that for every player i in N, every type $t^i \in T^i$ is of the form $t^i = (R^i, \tau^i)$, where R^i is the ordinal preference of t^i and τ^i represents the rest of the characteristics of t^i. This imposes the following constraints on the utility functions: $u^i(\cdot, (t^i, t^{-i}))$ must be a faithful representation of R^i where $t^i = (R^i, \tau^i)$. Thus, we are able to define Bayesian incentive compatibility in our model.

Definition 8 Consider the game $\Gamma_{Id} = (N; W, , \dots, W; T^1, \dots, T^n; p^1, \dots, p^n; u^1, \dots, u^n; d)$. The generalized decision scheme d is *Bayesian incentive compatible* (BIC) if truth telling is a BNE of the game Γ_{Id}. That is, the n-tuple of strategies $\tilde{s}_0 = (\tilde{s}_0^1, \dots, \tilde{s}_0^n)$, where $\tilde{s}_0^i(t^i) = (R^i, t^i)$ when $t^i = (R^i, \tau^i)$, for all $t^i \in T^i$ and $i \in N$, is a BNE of Γ_{Id}.

Unfortunately, there exist robust examples that possess no 'nice' (in a sense to be explained later) BIC solutions even in the complete information case as the following example shows.

Example 2 continued.

To Example 2 add the utility functions: For $\delta > 0$ and $\frac{2}{3}(1+\delta) < 1$ let

$$u^1(x) = u^2(z) = u^3(y) = 1 + \delta. \tag{21}$$

$$u^1(y) = u^2(x) = u^3(z) = 1. \tag{22}$$

$$u^1(z) = u^2(y) = u^3(x) = 0. \tag{23}$$

If we consider a decision scheme $d : W^N \to \Delta(A)$ to be a 'nice' representation of E if it satisfies the Condorcet condition[2](CC), then we claim that E has no BIC representation that satisfies CC. Assume, on the contrary that $d : W^N \to \Delta(A)$ is a representation of E satisfying CC and R^N is a Nash equilibrium of the game $G_d(N; W, W, W; u^1, u^2, u^3; d)$. Without loss of generality assume that $d(z; R^N) \geq \frac{1}{3}$. Then:

$$\sum_{a \in A} u^1(a) d(a; R^N) \leq \frac{2}{3}(1+\delta).$$

If player 1 deviates to $Q^1 = R^3$ then $d(y; Q^1, R^2, R^3) = 1$ since d is a representation of E that satisfies CC. Hence, this is a profitable deviation from truth telling since

$$\sum_{a \in A} u^1(a) d(a; R^3, R^2, R^3) = 1 > \frac{2}{3}(1+\delta).$$

Continuing our discussion of Example 2, we consider now the well-known Borda rule. This is an SCC in which the states in A are ranked by each player, in our case 2 (best), 1 (middle) or 0 (worst), and the chosen states are those with the maximal total score. For the profile of preferences in our example, each state scores 3 and hence all states are chosen according to the Borda rule. This is a representation of E (the simple majority rule) since no player is effective for any proper subset of $\{x, y, z\}$ and any two players can force any state, say x by submitting the preferences (x, y, z) and (x, z, y) (thus guaranteeing a score of at least 4 for x and at most 3 for each of y and z).

[2] A representation d of E satisfies the CC if for all $R^N \in W^N$, if $c \in A$ beats every $b \in A \setminus \{c\}$ by simple majority rule, then $d(c, R^N) = 1$.

The Borda rule does not satisfy the CC. To see this consider the profile of preferences:

$$R^N = \begin{array}{ccc} \underline{1} & \underline{2} & \underline{3} \\ x & x & y \\ y & y & z \\ z & z & x \end{array}$$

The Borda rule chooses $\{x, y\}$ although x is the Condorcet winner. Thus, the possibility that a decision scheme representing the Borda rule is a BIC is not excluded by the above proved claim. However, nevertheless, no decision scheme representing the Borda rule is BIC. To see this, consider the original profile R^N in our example:

$$R^N = \begin{array}{ccc} \underline{1} & \underline{2} & \underline{3} \\ x & z & y \\ y & x & z \\ z & y & x \end{array}$$

As we saw, the Borda rule selects $\{x, y, z\}$ and hence any decision scheme d representing it is of the form $d(x; R^N) = p_1,\ d(y; R^N) = p_2,\ d(z; R^N) = p_3$, where $p_1+p_2+p_3 = 1$. At leat one state is chosen with probability 1/3 say $d(z; R^N) \geq 1/3$. With the utility functions given in (21)–(23), the utility of player 1 is $p_1(1+\delta)+p_2$. By presenting the preference (y, x, z), player 1 guarantees utility 1 and this is a profitable deviation since:

$$p_1(1 + \delta) + p_2 < (p_1 + p_2)(1 + \delta) \leq \frac{2}{3}(1 + \delta) < 1.$$

Finally, we remark that there exist BIC representations of the effectivity function E in our example (which necessarily are not 'nice'). Let $d : W^N \to \Delta(A)$ satisfy $d(a; Q^N) = 1$ for all Q^N of the form:

$$Q^N = \begin{array}{ccc} \underline{S} & & \underline{N \setminus S} \\ a & a & \overline{Q^{N \setminus S}} \\ bc & bc & \end{array},$$

where $|S| = 2,\ \ a \in \{x, y, z\}$ and $\{b, c\} = \{x, y, z\} \setminus \{a\}$, and $d(\cdot; Q^N) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ otherwise. Then, d is a representation of E and the true preference profile R^N is a NE of the game $\Gamma = (N; W, W, W; u^1, u^2, u^3; d)$. This does not contradict the result of d'Aspremont and Peleg (1988) as their definition of representation is stronger than ours. Note also that this decision scheme is not 'nice': First, it clearly does not satisfy the CC; it does not choose (with certainty) the Condorcet winner x in the profile

$$R^N = \begin{array}{ccc} \underline{1} & \underline{2} & \underline{3} \\ x & x & y \\ y & z & z \\ z & y & x \end{array}$$

Second, it is not monotonic: By improving the position of z,

$$\text{from} \quad \begin{array}{ccc} \underline{1} & \underline{2} & \underline{3} \\ x & x & y \\ y & yz & z \\ z & & x \end{array} \quad \text{to} \quad \begin{array}{ccc} \underline{1} & \underline{2} & \underline{3} \\ x & x & y \\ yz & yz & z \\ & & x \end{array},$$

the probability of z decreases from 1/3 to 0.

2.5 Ex-post Pareto optimality of representations by decision schemes

We now investigate the possibility that our construction is Pareto optimal in some sense. First we need the following definition.

Definition 9 A generalized decision scheme $d : W^N \times T \to \Delta(A)$ is Pareto optimal ex-post if the following condition is satisfied:

$$[R^N \in W^N \text{ and } x \in A \text{ is not Pareto optimal w. r. t. } R^N] \Rightarrow d(x; R^N, t) = 0 \text{ for all } t.$$

It is possible to strengthen Theorem 4 by demanding that the solution d is also Pareto optimal ex-post. More precisely, the following result is true.

Theorem 6 *Let E be a monotonic and supperadditive effectivity function, let $\mathscr{I}$ be an information structure and let $u^1, \ldots, u^n$ be the utility functions of the players. Then E has a representation by a Pareto optimal ex-post generalized decision scheme d, such that the game $\Gamma_{Id} = (N; W, \ldots, W; \mathscr{I}; u^1, \ldots, u^n; d)$ has a BNE in pure strategies in which each player reports his true type.*

Proof We begin the proof of the theorem with some preliminary remarks. Let E be a monotonic and superadditive effectivity function. Then for every $R^N \in W^N$, the set $H(R^N) = PAR(R^N) \cap C_{uf}(E, R^N)$ is nonempty (here $PAR(R^N)$ is the set of Pareto optimal alternatives in A w.r.t. R^N). Indeed, if $x \in C_{uf}(E, R^N)$ and $y \in A$ satisfies yP^ix for all $i \in N$, then $y \in C_{uf}(E, R^N)$.

Our second claim is that $E^H = E$, which we deduce from Theorem 2 as follows: Since $H(R^N) \subseteq C_{uf}(R^N)$ for all $R^N \in W^N$, it follows from Theorem 2 that $E^H(S) \supseteq E(S)$ for all subsets $S \subseteq N$. To prove the converse inclusion, let $S \in P_0(N)$ and $B \in E^H(S)$. Then there exists $R^S \in W^S$ such that $H(R^S, Q^{N\setminus S}) \subseteq B$ for all $Q^{N\setminus S} \in W^{N\setminus S}$. In particular, $H(R^S, I^{N\setminus S}) \subseteq B$. By definition, $E^H(N) = E(N)$, and hence, we may assume that $S \neq N$. This implies that $PAR(R^S, I^{N\setminus S}) = A$. Therefore, $H(R^S, I^{N\setminus S}) = C_{uf}(R^S, I^{N\setminus S}) \subseteq B$. This implies, by Theorem 2, that $B \in E(S)$. In order to prove our theorem, it remains now to repeat the proof of Theorem 4 with $C_{uf}(E, R^N)$ replaced by $H(R^N)$. □

Remark that if the decision scheme of the last theorem is also BIC, then in BIC equilibrium, the final outcome is Pareto optimal.

2.6 Dichotomous preferences

In this subsection, we prove a variant of Theorem 4. For that, we first define a subset of W as follows:

Definition 10 A preference relation $R \in W$ is *dichotomous* if there exist $B_1, B_2 \in P(A)$ such that $B_1 \neq \emptyset$, $B_1 \cap B_2 = \emptyset$ and $B_1 \cup B_2 = A$ such that xIy if $x, y \in B_i$, $i = 1, 2$ and xPy if $x \in B_1$, $y \in B_2$. The set of all dichotomous preferences in W is denoted by W_δ.

Since a dichotomous preference relation is determined by a single subset $B \subseteq A$, the set of most preferred alternatives, we use the notation $R = \frac{B}{A \setminus B}$ for a generic dichotomous preference relation.

Lemma 1 *The social choice correspondence $H : W_\delta^N \to P_0(A)$ defined by $H(R^N) = C_{uf}(E, R^N)$ for all $R^N \in W_\delta^N$, is a representation of E.*

Proof We first prove the following claim: If $R^N \in W^N$, then there exists $R_1^N \in W_\delta^N$ such that $C_{uf}(E, R^N) = C_{uf}(E, R_1^N) = H(R_1^N)$. That is, for any profile of weak preferences on A, there exists a profile of dichotomous preferences with the same uniform core.

To see that, let $A \setminus C_{uf}(R^N) = \{x_1, \ldots, x_k\}$. By Abdou and Keiding (1991, p. 145), there exist disjoint coalitions $S_1, \ldots, S_k$ and sets $B_1, \ldots, B_k \in P_0(A)$such that for $j = 1, \ldots, k$, the outcome x_j is uniformly dominated by B_j via S_j at R^N. Define now R_1^N as follows:

- For $j = 1, \ldots, k$ and for $i \in S_j$; $xI_1^i y$ if $x, y \in B_j$ or $x, y \in A \setminus B_j$.
- For $j = 1, \ldots, k$ let $B_j P_1^{S_j} A \setminus B_j$.
- For $i \in N \setminus \cup_j S_j$ let $xI_1^i y$ for all $x, y \in A$.

It follows readily from the definition that $R_1^N \in W_\delta^N$ and that $C_{uf}(E, R^N) = C_{uf}(E, R_1^N) = H(R_1^N)$.

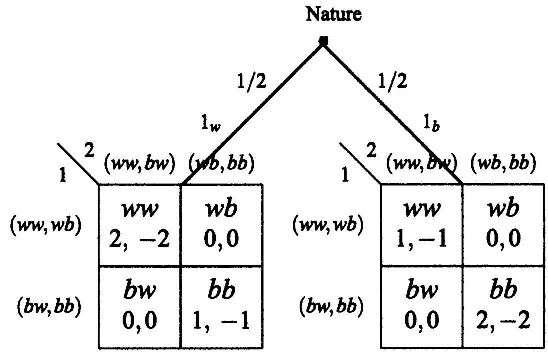
We now prove that $E^H = E^{C_{uf}}$. By Theorem 2, this will complete the proof of the lemma. Let $S \in P_0(N)$ and $B \in E^H(S)$. Then there exists $R^S \in W_\delta^S$ such that $H(R^S, Q^{N\setminus S}) \subseteq B$ for all $Q^{N\setminus S} \in W_\delta^{N\setminus S}$. In particular $H(R^S, I^{N\setminus S}) \subseteq B$ and, by Remark 3, $H(R^S, Q^{N\setminus S}) \subseteq B$ for all $Q^{N\setminus S} \in W^{N\setminus S}$, implying $B \in E^{C_{uf}}(S)$. Thus, $E^H(S) \subseteq E^{C_{uf}}(S)$ (in the usual set inclusion sense: $B \in E^H(S) \Rightarrow B \in E^{C_{uf}}(S)$), for all $S \in P_0(N)$.

In the other direction, let $S \in P_0(N)$ and $B \in E^{C_{uf}}(S)$. Then there exists $R^S \in W^S$ such that $C_{uf}(R^S, Q^{N\setminus S}) \subseteq B$ for all $Q^{N\setminus S} \in W^{N\setminus S}$, in particular $C_{uf}(R^S, I^{N\setminus S}) \subseteq B$. By the first step of the proof, there exists $R_1^S \in W_\delta^S$ such that

$$H(R_1^S, I^{N\setminus S}) = C_{uf}(R_1^S, I^{N\setminus S}) = C_{uf}(R^S, I^{N\setminus S}) \subseteq B.$$

By Remark 3 again, $H(R_1^S, Q^{N\setminus S}) = C_{uf}(R_1^S, Q^{N\setminus S}) \subseteq B$ for all $Q^{N\setminus S} \in W_\delta^{N\setminus S}$ implying that $B \in E^H(S)$ and hence $E^{C_{uf}}(S) \subseteq E^H(S)$. As this holds for all $S \in P_0(N)$, this completes the proof of Lemma 1. □

Fig. 1 The restricted game of Γ_δ



Using Lemma 1, we can repeat the proof of Theorem 4 to the game in which the players are restricted to dichotomous preference relations that is, replacing W by W_δ to obtain:

Theorem 7 *Let $E : P(N) \to P(P_0(A))$ be a monotonic and superadditive EF. Let $\mathcal{I} = (T^1, \dots, T^n; p^1, \dots, p^n)$ be an IS, and let $(u^1, \dots, u^n)$ be a vector of utilities for the players. Then E has a representation by a generalized decision scheme $d : W_\delta^N \times T \to \Delta(A)$ such that the game $\Gamma_\delta = (N; W_\delta, \dots, W_\delta; \mathcal{I}; (u^i)_{i\in N}; d)$ has a BNE in pure strategies in which each player reports his true type.*

2.7 Example 1 continued

Omitting the singleton type set of player 2 (and the trivial beliefs of player 1 on this type set), our information structure is $\mathcal{I} = (T^1, p^2)$ where $T^1 = \{1_w, 1_b\}$ and $p^2(1_w) = p^2(1_b) = 1/2$. We now define the utility functions of the agents:

- $u^1(ww, 1_w) = 2$, $u^1(bb, 1_w) = 1$ and $u^1(bw, 1_w) = u^1(wb, 1_w) = 0$ (1_w likes 'conformity' with preference to white shirts).
- $u^1(ww, 1_b) = 1$, $u^1(bb, 1_b) = 2$ and $u^1(bw, 1_b) = u^1(wb, 1_b) = 0$ (1_b likes 'conformity' with preference to blue shirts).
- $u^2(a, 1_w) = -u^1(a, 1_w)$ and $u^2(a, 1_b) = -u^1(a, 1_b)$ for all $a \in A$ (the utility of player 2 is 'opposed' to that of player 1 whatever his type is).

Consider the Bayesian game in which the players submit dichotomous preferences: $\Gamma_\delta = (N; W_\delta, W_\delta; \mathcal{I}; u^1, u^2; d_{uf})$. As a game in strategic form, this is a game in which player 2 has 15 pure strategies (indexed by the nonempty subsets of A) and player 1 has 15^2 pure strategies. In order to find a BNE, and hence a CE of this game, we focus on the following submatrix of Γ_δ described in Fig. 1 which we shall refer to as the 'restricted game.'

Here, the pure strategies are denoted by the upper-set in the dichotomous preference that is: $(ww, wb) \equiv \frac{ww,wb}{bw,bb}$ etc. Note that since player 1 is effective for the set $\{ww, wb\}$, simply by wearing a white shirt, playing the pure strategy (ww, wb) guarantees an outcome in $\{ww, wb\}$. Therefore, this strategy can be abbreviated as w (wearing a white shirt). Similarly for the other strategies in the reduced game. Thus,

Fig. 2 The restricted game in strategic form

1 \ 2	w	b
$(w,I_w),(w,I_b)$	$\frac{3}{2},-\frac{3}{2}$	$0,0$
$(w,I_w),(b,I_b)$	$1,-1$	$1,-1$
$(b,I_w),(w,I_b)$	$\frac{1}{2},-\frac{1}{2}$	$\frac{1}{2},-\frac{1}{2}$
$(b,I_w),(b,I_b)$	$0,0$	$\frac{3}{2},-\frac{3}{2}$

Fig. 3 A correlated equilibrium in the restricted game

1 \ 2	w	b
$(w,I_w),(w,I_b)$	0	0
$(w,I_w),(b,I_b)$	$\frac{2}{3}$	$\frac{1}{3}$
$(b,I_w),(w,I_b)$	0	0
$(b,I_w),(b,I_b)$	0	0

the reduced game is equivalent to the game with incomplete information on one side (on the side of player 2 regarding the type of player 1) in which the actions set of each player is $\{w, b\}$, wearing a white or a blue shirt.

A BNE of this restricted game is (s^1, s^2) where

$$s^1(1_w) = \frac{ww, wb}{bw, bb}, \quad s^1(1_b) = \frac{bw, bb}{bw, bb},$$

and

$$s^2 = \frac{1}{2}\,\frac{ww, bw}{wb, bb} + \frac{1}{2}\,\frac{wb, bb}{ww, bw}.$$

It can be shown that this is also a BNE of the game Γ_δ, and as far as we can see, Γ_δ has no BNE in pure strategies.

The strategic form (i.e., the ex-ante Harsanyi game) of the reduced game is thus given in Fig. 2. The strategies of player 1 are to be read in the natural way: (w, I_w), (b, I_b) means to play w when his type is I_w and play b when his type is I_b etc. A correlated equilibrium of this game is given in Fig. 3. The generalized decision scheme can now be defined by inserting this correlated equilibrium in Eqs. (12)–(14).

Concluding remarks

We have generalized in this paper the theory of representations of constitutions (see Peleg and Peters (2010, Part I)) to cover situations where the preferences of the citizens are private information. In our model, the constitution is specified by an effectivity function (like in Gardenfors 1981) and is common knowledge. Our representation is uniform in the sense that the set of actions of every player is dependent only on the set of social states and not on the effectivity function we are looking at. Actually, this set is the set of all weak orderings of the set of social states. We represent constitutions by means of generalized decision schemes, that is, functions from pairs of n-tuples of actions and types to lotteries on the set of social states. Thus, we follow a generalization of d'Aspremont and Peleg (1988). In our proofs, we rely heavily on the uniform core of the constitution. [The uniform core of an effectivity function was introduced in Abdou and Keiding (1991)].

Our main result, Theorem 4, is that for every superadditive and monotonic effectivity function and for every specification of types, beliefs, and utilities to the players, there exists a generalized decision scheme d such that: (1) d represents the effectivity function; and (2) the incomplete information game induced by d has a Bayesian Nash equilibrium in pure strategies. We have also checked possible extensions of the main result. The existence of Bayesian incentive compatible equilibria is not clear and may lead to pathological equilibria (Sect. 2.4). Ex-post Pareto optimality of (pure) equilibria may be obtained (Theorem 6). Finally, the common set of actions may be restricted to dichotomous preferences on the set of social states (Theorem 7).

Our results open the theory of representations of constitutions to unrestricted applications. In particular, the possibility to restrict our attention to dichotomous preferences looks promising. For an additional recent contribution to this area, see Peters et al. (H. Peters, M. Schröder and D. Vermeulen: Ex post Nash consistent representation of effectivity functions, 2013, unpublished).

Acknowledgments We are indebted to three anonymous referees for their helpful comments.

References

Abdou, J., Keiding, H.: Effectivity functions in social choice. Kluwer, Dordrecht (1991)
d'Aspremont, C., Peleg, B.: Ordinal Bayesian incentive compatible representations of committees. Soc. Choice Welf. **5**, 261–279 (1988)
Arrow, K.J.: Social Choice and Individual Values. Wiley, New York (1951, 1963)
Gardenfors, P.: Rights, games, and social choice. Nous **15**, 341–356 (1981)
Gibbard, A.: A Pareto-consistent libertarian claim. J. Econ. Theory **7**, 388–410 (1974)
Keiding, H., Peleg, B.: Binary effectivity rules. Rev. Econ. Des. **10**, 167–181 (2006)
Peleg, B.: Effectivity functions, game forms, games, and rights. Soc. Choice Welf. **15**, 67–80 (1998)
Peleg, B., Peters, H.: Strategic Social Choice. Springer, Berlin (2010)
Peleg, B., Peters, H., Storcken, T.: Nash consistent representation of constitutions: a reaction to the Gibbard Paradox. Math. Soc. Sci. **43**, 267–287 (2002)
Suzumura, K.: Welfarism, individual rights, and procedural fairness. In: Arrow, K.J., Sen, A., Suzumura, K. (eds.) Handbook of Social and Welfare, Volume 2, Chapter 23. North-Holland, Amsterdam (2011)