POWER INDICES AND THE VEIL OF IGNORANCE

Annick Laruelle and Federico Valenciano

WP-AD 2000-13

Correspondence to A. Laruelle: Universidad de Alicante, Campus de San Vicente, E-03071 Alicante, Spain. (e-mail: laruelle@merlin.fae.ua.es).

Editor: Instituto Valenciano de Investigaciones Económicas, s.a.

First Edition June 2000.

Depósito Legal: V-2092-2000

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

---

* We want to thank M. Maschler and J. M. Zarzuelo for their insightful comments. This research has been supported by the Training and Mobility of Researchers programme initiated by the European Commission, by the IVIE, and by the DGES of the Spanish Ministerio de Educación y Cultura, under project PB96-0247.

* A. Laruelle: Universidad de Alicante. F. Valenciano: Universidad del País Vasco.

POWER INDICES AND THE VEIL OF IGNORANCE

Annick Laruelle and Federico Valenciano

Abstract

We provide an axiomatic foundation of the expected utility preferences over lotteries on roles in simple superadditive games represented by the two main power indices, the Shapley-Shubik index and the Banzhaf index, when they are interpreted as von Neumann-Morgenstern utility functions. Our axioms admit meaningful interpretations in the setting proposed by Roth in terms of di®erent attitudes toward risk involving roles in collective decision procedures under the veil of ignorance. In particular, an illuminating interpretation of "e±ciency", up to now missing in this set up, as well as of the corresponding axiom for the Banzhaf index, is provided.

Key words: Power indices, voting power, collective decision-making, lotteries, expected utility

1 INTRODUCTION

In the literature, decision-making procedures are often modeled as simple games by assigning worth 1 to coalitions with the capacity of passing a decision and 0 to the others. In this framework, measures have been developed in order to assess the a priori distribution of power among the players, that is, their capacity to in°uence the outcome of a vote. The two best known power indices are the Shapley-Shubik (1954) index and the Banzhaf (1965) index. Both indices can be interpreted either as a measure axiomatically grounded or as a probability. In the ¯rst approach, each power index is interpreted as the unique measure embodying a set of properties that characterize it. Since Dubey's (1975) ¯rst axiomatization of the Shapley-Shubik index on the domain of simple games and that of Dubey and Shapley (1979) of the Banzhaf index, several axiomatization have been proposed in the literature. However, most axiomatizations of these indices pay little attention to the compellingness or even to the meaning of the axioms in terms of the voting situations underlying simple games. Alternatively, in the second approach, the value of both indices for a player in a game is interpreted as the a priori probability of that player being a swinger in the coalition passing a decision that is made according to the voting rules modeled by that game. Either index, as any semivalue (Weber (1988), Einy (1987)), emerges then from di®erent probabilistic assumptions about this coalition (see also Stra±n (1977)). In fact, the axiomatic and the probabilistic approaches are complementary interpretations that do not ¯t clearly with each other, as pointed out by Blair and McLean (1990), who propose an axiomatic foundation for the probabilistic approach.

In this paper we consider a third interpretation, that was proposed by Roth (1977b, see also 1988), without attracting so far much attention in the literature. In this approach power indices are interpreted as utility functions representing von Neumann-Morgenstern preferences over the set of lotteries on positions or roles in collective decision-making procedures. In other words, the value that an index attaches to position i in game v is just a means to allow comparisons of the capacity to in°uence the outcome in position i in a vote cast according the rules described by v; with the capacity attached to other positions in other voting rules or even random mixtures of them. Under this point of view, what matters of the information provided by an index are the von Neumann-Morgenstern preferences it represents. Consequently, in this sense an index is equivalent to any positive a±ne transformation of it. In Laruelle and Valenciano (1999) we provide an axiomatization of both power indices up to the choice of a zero and a unit of scale, that is, exactly the natural degrees of indeterminacy for a von Neumann-Morgenstern utility function. It turns out that all our axioms in (1999) admit a direct translation into Roth's setting. As a result an axiomatization in this setting emerges in which our axioms acquire a clear interpretation in terms of attitudes toward risk involving roles in collective decision procedures. In particular, several alternatives to Roth's "neutrality to ordinary risk", a direct translation of Dubey's (1975) "transfer", are provided. While his obscure "neutrality/aversion to strategic risk" that differentiates either index is replaced by an illuminating interpretation of "efficiency", up to now missing in this set up, as well as of the corresponding axiom for the Banzhaf index, in terms of attitudes toward risk involving roles in collective decision procedures under the veil of ignorance (Rawls, 1972) with respect to the role to play.

The paper is organized as follows. In Section 2 the basic game theoretical background is given along with a summary of our characterization of the Shapley-Shubik and Banzhaf indices in Laruelle and Valenciano (1999). In Section 3 Roth's (1977b) setting and characterization are briefly reviewed. In Section 4 we translate the axioms reviewed in Section 2 to Roth's setting, where a new characterization is provided. Finally Section 5 concludes with a brief discussion emphasizing the main conclusions of this work and some lines for further research.

2 SHAPLEY-SHUBIK AND BANZHAF INDICES UP A ZERO AND A UNIT

A cooperative transferable utility (TU) game is a pair $(N; v)$, where $N = \{1; \ldots; n\}$ denotes the set of players and v a function which assigns a real number to each non-empty subset or coalition of N, and $v(\emptyset) = 0$. When N is clear from the context, we will refer to game $(N; v)$ as game v. The number of players in a coalition S is denoted s. A game is monotonic if $v(T) \geq v(S)$ whenever $T \supseteq S$. It is superadditive if $v(S \cup T) \geq v(S) + v(T)$ whenever $S \cap T = \emptyset$.

A (0-1)-game is a game in which the function v only takes the values 0 and 1. It is a simple game if it is not identically 0; and monotonic. In this context, the superadditivity property is equivalent to the condition: $v(S) + v(N \setminus S) \leq 1$ for all $S \subset N$. Let SG_n denote the set of all simple superadditive n-person games. The following definitions refer to games in SG_n. A coalition S is said to be winning in game v if $v(S) = 1$, and is losing if $v(S) = 0$. A winning coalition is minimal if it does not contain any other. In game v; $W(v)$ (resp., $M(v)$) will denote the set of winning (resp., minimal winning) coalitions in v. Any of these sets, $W(v)$ or $M(v)$, fully characterizes the game v. A player i is said to be a swinger in a coalition S if S is winning and $S \setminus \{i\}$ is not. A null player in a game v is a player i who is never a swinger, that is, $v(S) = v(S \setminus \{i\})$ for all S.

As a collective decision-making procedure is specified by the voting body and the decision rules, it can be modeled by a (0-1)-game whose winning coalitions are those that can make a decision without the vote of the remaining players. We assume that the decision rules are consistent in the following sense. The unanimity of the players can make a decision. Any subgroup of a group of voters that cannot make a decision cannot either. Two nonintersecting groups of voters cannot make decision at the same time. Under these conditions a voting procedure can be described as a simple superadditive game. Notwithstanding, note that such a voting procedure is fully specified by the list of winning coalitions, and any assumption concerning the meaning of the numerical 0/1 values of the game is unnecessary and unjustified. In particular the assumption, common in the TU context, that the value of a coalition S represents the utility that the members of S can distribute among themselves is out of place here.

For any coalition $S \subseteq N$, the S-unanimity game, denoted $(N; u^S)$, is the simple game

$$u^S(T) = \begin{cases} 1 & \text{if } T \supseteq S \\ 0 & \text{otherwise.} \end{cases}$$

Player i's dictatorship is thus denoted by $u^{\{i\}}$. For any game $v \in SG_n$ such that $v \neq u^N$, and any $S \in M(v)$, the modified game v_S^* is the game such that $W(v_S^*) = W(v) \setminus \{S\}$. Avoiding starting from the unanimity game and dropping a minimal winning coalition guarantee that $v_S^* \in SG_n$. In terms of decision-making procedures, the modified game v_S^* represents the new procedure that results from the modification of a decision-making rules in such a way that one and only one coalition that previously could make a decision cannot any more.

A power index is a function $\Phi : SG_n \to \mathbb{R}^n$ that associates with each simple superadditive game v a vector or power profile $\Phi(v)$ whose ith component is interpreted as an a priori measure of the influence that player i can exert on the outcome when decisions are to be made according to the decision rule described by v. To evaluate the distribution of power among the players the two best known power indices are the Shapley-Shubik (1954) index and the Banzhaf (1965) index. The Shapley-Shubik index is given by

$$Sh_i(v) = \sum_{\substack{S \subseteq N \\ (S \ni i)}} \frac{(s-1)!(n-s)!}{n!} (v(S) - v(S \setminus \{i\})); \quad i = 1, \ldots, n.$$

The Banzhaf index is given by

$$Bz_i(v) = \frac{1}{2^{n-1}} \sum_{\substack{S \subseteq N \\ (S \ni i)}} (v(S) - v(S \setminus \{i\})); \quad i = 1, \ldots, n.$$

The Shapley-Shubik index was first axiomatized by Dubey (1975). A similar axiomatization of the Banzhaf index sharing three out of four axioms was given by Dubey and Shapley (1979).

The main purpose of these or any other measure of power is to allow comparisons of the players' capacity to influence the outcome in the same or in different voting procedures. Moreover, as already pointed out above, the only relevant information in a decision-making procedure is the list of winning coalitions. Therefore a power index should be based on this information, and should not attribute any importance to the particular numerical 0/1 values of the characteristic function used to describe it. These considerations took us to propose in Laruelle and Valenciano (1999) a characterization consistent with this idea. More precisely, avoiding any normalizing ingredients in our axioms we characterized both indices up to the choice of a zero and a unit. We now summarize our axiomatization, first reviewing our axioms, for we only share anonymity with previous characterizations in the literature.

Anonymity (An): For all $v \in SG_n$, any permutation π of N, and any $i \in N$,

$$\Phi_i(\pi v) = \Phi_{\pi(i)}(v),$$

where $(\pi v)(S) := v(\pi(S))$.

This axiom states that the measure of power does not depend on how the players are labeled.

Null Player (NP): For all $v \in SG_n$, and all $i \in N$,

$$i \text{ is a null player in } v \Leftrightarrow \text{ for all } w \in SG_n, \ \Phi_i(v) \leq \Phi_i(w).$$

The axiom states that being a null player is the (strictly, mind the equivalence) worst role any player can play, the role that yields a minimal measure of power.

Transfer (T): For any $v, w \in SG_n$, and all $S \in M(v) \cap M(w)$ $(S \neq N)$:

$$\Phi_i(v) - \Phi_i(v_S^*) = \Phi_i(w) - \Phi_i(w_S^*) \quad (\forall i \in N).$$

This axiom, equivalent to the usual transfer, postulates that the effect (gain or loss) on any player's power of eliminating a single minimal winning coalition from the set of winning ones is the same in any game in which this coalition is minimal winning. In Laruelle and Valenciano (1999) we also show that this axiom can be replaced in the characterizing theorem by the following weaker (under anonymity) assumption.

Symmetric Gain-Loss (SymGL): For all $v \in SG_n$, all $S \in M(v)$ $(S \neq N)$, and all $i, j \in S$ (resp., $i, j \in N \setminus S$), $\Phi_i(v) - \Phi_i(v_S^*) = \Phi_j(v) - \Phi_j(v_S^*)$.

The axiom states that the effect of eliminating a minimal winning coalition is the same for any two players belonging to it and also for any two players outside it. Finally, our differentiating axioms for the Shapley-Shubik and the Banzhaf index are respectively:

Total Gain-Loss Balance (TGLB): For all $v \in SG_n$ and all $S \in M(v)$ $(S \neq N)$,

$$\sum_{i \in S} (\Phi_i(v) - \Phi_i(v_S^*)) = \sum_{j \in N \setminus S} (\Phi_j(v_S^*) - \Phi_j(v)).$$

Average Gain-Loss Balance (AGLB): For all $v \in SG_n$ and all $S \in M(v)$ $(S \neq N)$,

$$\frac{1}{s} \sum_{i \in S} (\Phi_i(v) - \Phi_i(v_S^*)) = \frac{1}{n-s} \sum_{j \in N \setminus S} (\Phi_j(v_S^*) - \Phi_j(v)).$$

Total (resp., average) gain-loss balance postulates that the total (resp., average) loss of the players in a minimal winning coalition equals the total (resp., average) gain of the players outside it, when this coalition is eliminated from the list of winning coalitions.

Now, denoting $\mathbf{1} := (1, ..., 1) \in \mathbb{R}^n$, we have the main result in Laruelle and Valenciano (1999): our axioms characterize both indices are up to the choice of a zero and a unit of scale for the measure of power.

Theorem 1 (Laruelle and Valenciano, 1999) Let $\Phi : SG_n \to \mathbb{R}^n$, then

(i) Φ satisfies anonymity, null player, symmetric gain-loss (or transfer) and total gain-loss balance if and only if it is $\Phi = \alpha Sh + \beta \mathbf{1}$, for some $\alpha > 0$ and $\beta \in \mathbb{R}$.

(ii) Φ satisfies anonymity, null player, symmetric gain-loss (or transfer) and average gain-loss balance if and only if it is $\Phi = \alpha Bz + \beta \mathbf{1}$, for some $\alpha > 0$ and $\beta \in \mathbb{R}$.

3 ROTH'S SETTING AND RESULTS

Let $L(SG_n \times N)$ denote the set of lotteries on $SG_n \times N$. That is, the set of distributions of probability on the set $SG_n \times N$[1]. A pair (v, i) in this set is interpreted as the prospect of the event "playing role i in game v". Any power index $\Phi : SG_n \to \mathbb{R}^n$ determines a preference ordering on $L(SG_n \times N)$, the one represented by the expected utility function associated to $\varphi : SG_n \times N \to \mathbb{R}$, defined by $\varphi(v, i) := \Phi_i(v)$. Any such an ordering will rank all possible roles in all possible games, as well as lotteries on them. This is the setting proposed by Roth (1977b, 1988), in which the Shapley-Shubik and Banzhaf indices are reinterpreted as utility functions representing risk preferences. This interpretation implies a point of view beyond that of any particular player. There are no players in fact, just

[1] We assume $SG_n \times N \subset L(SG_n \times N)$ by identifying at all effects the sure event (v, i) and the lottery that assigns probability one to this event.

"positions" in games or "roles", hence the more appropriate expression in this context "playing role i in game v". In this setting Roth investigates the attitudes toward risk that underlie and characterize the preferences associated to either index. But before proceeding with Roth's results we need some notation.

A lottery on $SG_n \times N$ can be represented by a map $l : SG_n \times N \rightarrow \mathbb{R}$, such that (i) for all (v, i) in $SG_n \times N$, $l(v, i) \geq 0$, and (ii) $\sum_{i \in N} \sum_{v \in SG_n} l(v, i) = 1$, where $l(v, i)$ is the probability of playing role i in game v. Given $l, l' \in L(SG_n \times N)$, and $\lambda \in [0, 1]$, $\lambda l \oplus (1 - \lambda) l'$ will denote the lottery such that $(\lambda l \oplus (1 - \lambda) l')(v, i) := \lambda l(v, i) + (1 - \lambda) l'(v, i)$. That is, $\lambda l \oplus (1 - \lambda) l'$ can be interpreted as a second order lottery that assigns probabilities λ and $1 - \lambda$ to l and l' respectively. Any "convex combination" of lotteries can be defined similarly. With this notation, for instance, $\frac{1}{2}(v, i) \oplus \frac{1}{2}(w, j)$ will denote the lottery that gives one half to playing role i in game v and one half to playing role j in game w. The support of a lottery l is the set $sup(l) := \{(v, i) \in SG_n \times N \text{ s.t. } l(v, i) > 0\}$. Two special kinds of lotteries will play an important role. On the one hand, lotteries in which the game to be played is sure but the role is random. In particular $\oplus_{i \in N} \frac{1}{n}(v, i)$ will denote the lottery that assigns the same probability $\frac{1}{n}$ to all roles in game v. Also, for any $S \subset N$, we will write $\oplus_{i \in S} \frac{1}{s}(v, i)$ with obvious similar meaning. On the other hand, lotteries in which the game to be played is random but the role is sure will be used too. Such lotteries can be expressed consistently with the former notation like this $\oplus_{v \in SG_n} l(v, i)(v, i)$, if role i is sure.

We now formulate Roth's assumptions relative to an ordering $\preceq$ in $L(SG_n \times N)$ that permitted him to single out the specific ones associated to Shapley-Shubik and Banzhaf indices. The indifference relation associated to $\preceq$ will be denoted "$\sim$".

Definition 1 *A binary relation $\preceq$ is a von Neumann-Morgenstern (VNM) preference[2] ordering on $L(SG_n \times N)$, if there exists a map $\varphi : SG_n \times N \rightarrow \mathbb{R}$, whose associated expectation $\mu : L(SG_n \times N) \rightarrow \mathbb{R}$, given by $\mu(l) := \sum_{i \in N} \sum_{v \in SG_n} l(v, i) \varphi(v, i)$, represents $\preceq$; that is, $l \preceq l'$ if and only if $\mu(l) \leq \mu(l')$.*

R1: For all $(v, i) \in SG_n \times N$, any permutation π of N, and any $i \in N$,

$$(\pi v, i) \sim (v, \pi i).$$

R2: For all $(v, i) \in SG_n \times N$, if v_0 denotes the 0-game s.t. $v_0(S) = 0$ for all S,

(i) $(v_0, i) \preceq (v, i) \preceq (u^{\{i\}}, i)$.

(ii) i is a null player in $v \Rightarrow (v_0, i) \sim (v, i) \prec (u^{\{i\}}, i)$.

[2]See von Neumann and Morgenstern (1944, 1947, 1953) and also Herstein and Milnor (1953) for its axiomatic foundation.

Neutrality to ordinary risk: For all $v, w \in SG_n$, and all $i \in N$,

$$\tfrac{1}{2}(v, i) \oplus \tfrac{1}{2}(w, i) \sim \tfrac{1}{2}(v \vee w, i) \oplus \tfrac{1}{2}(v \wedge w, i).$$

Neutrality to strategic risk: For any $S \subseteq N$, and all $i \in S$,

$$(u^S, i) \sim \tfrac{1}{s}(u^{\{i\}}, i) \oplus (1 - \tfrac{1}{s})(v_0, i).$$

Banzhaf-aversion to strategic risk: For any $S \subseteq N$, and all $i \in S$,

$$(u^S, i) \sim \tfrac{1}{2^{s-1}}(u^{\{i\}}, i) \oplus (1 - \tfrac{1}{2^{s-1}})(v_0, i).$$

Then Roth's characterizations can be restated like this omitting the straightforward normalizing requirements to get precisely Sh and Bz[3]:

Theorem 2 (Roth 1977b, 1988) The only von Neumann-Morgenstern preference ordering on $L((SG_n \cup \{v_0\}) \times N)$ that satisfies conditions R1, R2, neutrality to ordinary risk and neutrality to strategic risk (resp., Banzhaf-aversion to strategic risk) is the one represented by the utility function $sh(v, i) := Sh_i(v)$ (resp., $bz(v, i) := Bz_i(v)$).

Some remarks are worth here. First, note that "R1" corresponds to the traditional "anonymity" in the usual set up, while "neutrality to ordinary risk" is the direct translation of Dubey's transfer axiom. As to condition "R2", that unnecessarily uses the nonsimple zero-game v_0, is an awkward translation of traditional "null player" including some additional plausible ingredients. The role of the differentiating axioms is played by "neutrality to strategic risk" and a form of "aversion to strategic risk". But the translation of these two axioms from the preferences setting back into usual set up is the following. The first one's counterpart would be just assuming that for any S-unanimity game $\Phi_i(u^S) = \frac{1}{s} = Sh_i(u^S)$. As to the second, its counterpart would be just assuming that for any S-unanimity game $\Phi_i(u^S) = \frac{1}{2^{s-1}} = Bz_i(u^S)$. This is just imposing the "right" value of the index for the unanimity games in either case. In other words, the counterparts of these axioms in the usual set up are rather ad hoc assumptions that would permit to derive directly either index

[3] In Roth (1977a) he first reinterprets and axiomatizes the Shapley value as a von Neumann-Morgenstern utility function representing preferences over lotteries on positions in superadditive TU games. Then, in (1977b), he adapts his axiomatization for recasting Dubey's (1975) characterization of the Shapley-Shubik index into the domain of lotteries on positions in simple superadditive games, and also characterizes in this setting both the "raw" ("nonnormalized" in his terms) and normalized Banzhaf in this domain. Finally, in (1988), he presents a synthesis integrating both papers in which also characterizes the Banzhaf semivalue in this framework and the characterizations are adapted to the case of simple games. Although this only done for the Shapley-Shubik index, a similar adaptation for the Banzhaf semivalue is straightforward.

by just assuming, in addition, null player and transfer, so making superfluous anonymity and efficiency or the corresponding differentiating axiom for Banzhaf. Moreover, Roth's interpretation of these axioms on "strategic" grounds, however appealing at first sight, are particularly misleading in his set up. In either case, the condition and its game-theoretic flavored name, is explained in the following terms. Both axioms postulate the indifference between a certain gamble involving playing the dictator role or playing the zero-game, and the sure membership of the unique minimal winning coalition in the S-unanimity game u^S. Roth's interpretation is that playing the game u^S involves a certain "strategic" (rather than probabilistic, for no gamble is involved) risk. Thus: "..the two indices reflect different attitudes toward the relative benefits of engaging in strategic interaction with other players in games of the form u^S:" This game-theoretical explanation does not make sense in a context in which no strategic consideration, nor even players are involved. We find it contradictory with the most interesting and illuminating interpretation of Roth's setting.

4 BEHIND THE VEIL OF IGNORANCE

In our view, the only consistent interpretation of Roth's setting is that what matters is the preference ordering on roles in voting rules when one is uncertain with respect to the role to be played. That is, in the more suitable and suggestive Rawls' (1972) terms, "under the veil of ignorance". It is in these terms that all axioms should be interpreted. Then the following assumptions are the result of translating into the present framework the axioms reviewed in Section 2, used by us in Laruelle and Valenciano (1999) to characterize the Shapley-Shubik and Banzhaf indices up to a zero and a unit in SG_n: Namely, in what follows all assumptions refer to a preference ordering $\succsim$ on $L(SG_n \times N)$ (for we will not use the zero-game). In order to make it clearer this correspondence we use the same names, full or abbreviated, just adding one asterisk.

Anonymity* (An*): For all $(v, i) \in SG_n \times N$, any permutation π of N, and any $i \in N$,

$$(\pi v, i) \sim (v, \pi i),$$

where $(\pi v)(S) := v(\pi(S))$.

This is just Roth's R1, the only axiom common to his characterization and ours.

Null Player* (NP*): For all $(v, i) \in SG_n \times N$,

$$i \text{ is a null player in } v \Leftrightarrow \text{for all } w \in SG_n, \ (v, i) \precsim (w, i).$$

This axiom is the direct translation of our null player (NP): The role of null player is the worst (strictly) that can be attached to any position i.

Transfer* (T*): For any $v, w \in SG_n$; all $S \in M(v) \cap M(w)$ $(S \neq N)$; and all $i \in N$

$$\frac{1}{2}(v,i) \oplus \frac{1}{2}(w^*_S,i) \sim \frac{1}{2}(v^*_S,i) \oplus \frac{1}{2}(w,i).$$

This axiom postulates the indifference, whenever the position i is sure and S is a minimal winning coalition of games v and w, between the lottery that gives identical probabilities to play v or w^*_S and the lottery that gives identical probability to play v^*_S or w. Note that this axiom, a direct translation of our transfer (T), is equivalent to Roth's (1977b, 1988) "neutrality to ordinary risk" though the involved games are simpler. Observe that for any reasonable preference, if $i \in S$ (if $i \in N \setminus S$ the preferences should be reversed); $(v,i) \succ (v^*_S,i)$ and $(w^*_S,i) \prec (w,i)$, then the assumption expresses the intensity of these desirability comparisons.

As we will show, also in this framework transfer* can be replaced by the following simpler (and weaker under anonymity*) condition that results from translating our symmetric gain-loss into this setting.

Symmetric Gain-Loss* (SymGL*): For all $v \in SG_n$, all $S \in M(v)$ $(S \neq N)$, and all $i, j \in S$ (resp., $i, j \in N \setminus S$),

$$\frac{1}{2}(v,i) \oplus \frac{1}{2}(v^*_S,j) \sim \frac{1}{2}(v^*_S,i) \oplus \frac{1}{2}(v,j).$$

The axiom postulates the indifference between the lottery that gives identical probability to role i in v and to role j in v^*_S and the lottery that gives identical probability to role j in v and to role i in v^*_S, given that both players i and j are either both in the minimal winning coalition S dropped or both outside it. Now, as for any reasonable preference, if $i, j \in S$ (if $i, j \in N \setminus S$ the preferences should be reversed); $(v,i) \succ (v^*_S,i)$ and $(v^*_S,j) \prec (v,j)$, again the assumption expresses the intensity of these desirabilities.

As an alternative to transfer* or symmetric gain-loss*, the following axiom, that is not the translation of any of the axioms reviewed in Section 2 and is stronger than transfer* as shown in Proposition 1, has a clear and compelling interpretation. Thus, it can either replace or justify any of the two former axioms in the characterizing theorem.

Coalitional Expectations Dependence (CED): For all $l, l' \in L(SG_n \times N)$ with support in $SG_n \times \{i\}$ for some $i \in N$,

$$\sum_{v \in SG_n} l(v,i)v = \sum_{v \in SG_n} l'(v,i)v \quad \Rightarrow \quad l \sim l'.$$

This axiom requires that the ranking of lotteries in which the position is sure depends exclusively on the coalitional expectations of being winning. In other words, two lotteries in which the same position is sure and that assign to each coalition the same probability of being winning should be considered indifferent. Dubey and Shapley (1979) point out that this property, though they do not state it in general terms, would justify transfer. Also Roth (1977b) alludes to this property being satisfied by the lotteries involved in his neutrality to ordinary risk. Also note that the left-hand side of the implication is an equality on games in the convex hull of simple superadditive games not on lotteries[4]. This equality holds if and only if in the lotteries on SG_n, $\oplus_{v \in SG_n} l(v;i)v$ and $\oplus_{v \in SG_n} l'(v;i)v$ the expectation of any coalition of being winning is the same. The axiom postulates the indifference of l and l' in such a case. This condition seems quite natural. In the usual domain SG_n power indices rank games, each of them consisting of a list of winning coalitions. Now, when the role is fixed, to each lottery on SG_n is associated a list in which each coalition is winning with some probability. Coalitional expectation dependence just requires that this list is what determines the ranking of a lottery.

The following propositions establish the relationships between the former axioms.

Proposition 1 Coalitional expectations dependence (CED) implies transfer* (T*).

Proof. Let $v, w \in SG_n$, and $S \in M(v) \cap M(w)$ ($S \neq N$). Just observe that for all $T \subseteq N$, $\frac{1}{2}v(T) + \frac{1}{2}w_S^*(T) = \frac{1}{2}v_S^*(T) + \frac{1}{2}w(T)$. Indeed, if $T \neq S$: $v(T) = v_S^*(T)$ and $w(T) = w_S^*(T)$; and if $T = S$: $v(T) = w(T) = 1$ and $v_S^*(T) = w_S^*(T) = 0$. Thus, by CED, $\frac{1}{2}(v,i) \oplus \frac{1}{2}(w_S^*,i) \sim \frac{1}{2}(v_S^*,i) \oplus \frac{1}{2}(w,i)$ for all $i \in N$. ■

Proposition 2 Any von Neumann-Morgenstern (VNM) ordering satisfying anonymity* (An*) and transfer* (T*), satisfies symmetric gain-loss* (SymGL*).

Proof. Let $v \in SG_n$, and $S \in M(v)$ ($S \neq N$). First note that for any permutation π of N, $(\pi v)_S^* = \pi(v_{\pi S}^*)$. Now let $i, j \in S$, and let π the permutation interchanging i and j. Then $S = \pi S$, so that $S \in M(v) \cap M(\pi v)$ and $(\pi v)_S^* = \pi(v_{\pi S}^*) = \pi(v_S^*)$. By VNM, An* and T*, $\frac{1}{2}(v,i) \oplus \frac{1}{2}(v_S^*,j) = \frac{1}{2}(v,i) \oplus \frac{1}{2}(v_S^*,\pi i) \sim \frac{1}{2}(v,i) \oplus \frac{1}{2}(\pi(v_S^*),i) = \frac{1}{2}(v,i) \oplus \frac{1}{2}((\pi v)_S^*,i) \sim \frac{1}{2}(v_S^*,i) \oplus \frac{1}{2}(\pi v,i) \sim \frac{1}{2}(v_S^*,i) \oplus \frac{1}{2}(v,\pi i) = \frac{1}{2}(v_S^*,i) \oplus \frac{1}{2}(v,j)$. The case $i, j \in N \setminus S$ is entirely similar. ■

Observe that only through the implied "sustitutivity" the VNM assumption has played a role in the proof of the previous proposition. In Laruelle and Valenciano (1999) an

[4] In fact, this axiom is an explicit statement of the assumption underlying our identification of $L(SG_n)$ and $Co(SG_n)$ in Laruelle and Valenciano (1998).

example shows that under anonymity transfer is strictly stronger than symmetric gain-loss. The same example can be adapted to show that the converse of Proposition 2 for T* and SymGL* is not true.

Finally the following axioms, interpretable as different forms of indifference through the veil of ignorance, are the translations into this setting of our differentiating axioms, total and average gain-loss balance.

Absolute Indifference under the Veil of Ignorance (AIVI): For all $v \in SG_n$, and all $S \in M(v)$ $(S \neq N)$,

$$\bigoplus_{i\in N} \frac{1}{n}(v;i) \sim \bigoplus_{i\in N} \frac{1}{n}(v_S^*;i):$$

This axiom postulates the indifference between playing v or v_S^* when all positions are equally probable. In fact, it can be easily seen that it is equivalent to stating the indifference of playing any two games when all roles are equally probable. For it, just note that whatever the game v, by dropping minimal winning coalitions one each time the unanimity game u^N is finally reached. So, by repeatedly applying the axiom it follows that when all roles are equally probable v and u^N, and consequently any two games, are indifferent. This is in fact the natural counterpart of "efficiency" once stripped of its normalizing ingredients and translated to Roth's setting.

Conditional Indifference under the Veil of Ignorance (CIVI): For all $v \in SG_n$, and all $S \in M(v)$ $(S \neq N)$,

$$\frac{1}{2}\Big(\bigoplus_{i\in S} \frac{1}{s}(v;i)\Big) \oplus \frac{1}{2}\Big(\bigoplus_{i\in N\setminus S} \frac{1}{n-s}(v;i)\Big) \sim \frac{1}{2}\Big(\bigoplus_{i\in S} \frac{1}{s}(v_S^*;i)\Big) \oplus \frac{1}{2}\Big(\bigoplus_{i\in N\setminus S} \frac{1}{n-s}(v_S^*;i)\Big):$$

This axiom postulates the indifference between playing v or v_S^* when it is equally probable to play a role in S or in $N \setminus S$, and all roles are equally probable within each of these coalitions. In other words, now it is indifferent to play v or v_S^* if the way of assigning the role is the following: first, a coin is tossed to choose S or $N \setminus S$, then a role is chosen at random in the previously chosen coalition.

To make more clear a comparison between these two different attitudes facing the veil of ignorance, observe that both conditions are particular cases of the following principle:

For all $v \in SG_n$, and all $S \in M(v)$ $(S \neq N)$,

$$\lambda_s\Big(\bigoplus_{i\in S} \frac{1}{s}(v;i)\Big) \oplus (1-\lambda_s)\Big(\bigoplus_{i\in N\setminus S} \frac{1}{n-s}(v;i)\Big) \sim \lambda_s\Big(\bigoplus_{i\in S} \frac{1}{s}(v_S^*;i)\Big) \oplus (1-\lambda_s)\Big(\bigoplus_{i\in N\setminus S} \frac{1}{n-s}(v_S^*;i)\Big);$$

for some collection $\lambda = (\lambda_s)_{s=1,2,\ldots,n-1}$, with $\lambda_s \in (0,1)$:

AIVI is the particular case $\lambda_s = \frac{s}{n}$; while CIVI is the particular case $\lambda_s = \frac{1}{2}$. That is, assuming all roles in S and $N \setminus S$ are equally probable, the indifference between v

and v_S^* depends on the way of choosing between S and $N \setminus S$. Ceteris paribus, from CIVI's point of view what matters is to be or not to be in S. While according to AIVI the size of the coalition matters too. More precisely, according to AIVI it is as if all the advantage of being or not in S or $N \setminus S$ were to be finally assigned to only one player in the coalition chosen at random. So that the importance of being in the right coalition has to be inversely weighted by its size.

Before proceeding with the main result, let us see that the former axioms, except coalitional expectations dependence, are one by one the exact translation into Roth's setting of the axioms reviewed in Section 2.

Proposition 3 Let $\Psi : SG_n \to \mathbb{R}^n$ and $\varphi : SG_n \times N \to \mathbb{R}$, such that $\Psi_i(v) = \varphi(v; i)$ for all $(v; i) \in SG_n \times N$. Then Ψ satisfies any of the following conditions: An, NP, SymGL, T, TGLB or AGLB, if and only if μ_φ satisfies the corresponding condition, that is, An*, NP*, SymGL*, T*, AIVI or CIVI, respectively.

Proof. Let $\Psi_i(v) = \varphi(v; i)$ for all $(v; i) \in SG_n \times N$. First, the following two equivalences are straightforward.

(An/An*): Ψ satisfies An $\Leftrightarrow$ μ_φ satisfies An*.

(NP/NP*): Ψ satisfies NP $\Leftrightarrow$ μ_φ satisfies NP*.

Now let us check the others one by one.

(SymGL/SymGL*): Let $v \in SG_n$, $S \in M(v)$ $(S \neq N)$, and $i, j \in S$ (resp., $i, j \in N \setminus S$). Then the following equivalences are immediate:

$$\Psi_i(v) - \Psi_i(v_S^*) = \Psi_j(v) - \Psi_j(v_S^*) \Leftrightarrow \Psi_i(v) + \Psi_j(v_S^*) = \Psi_j(v) + \Psi_i(v_S^*)$$

$$\Leftrightarrow \frac{1}{2}\varphi(v; i) + \frac{1}{2}\varphi(v_S^*; j) = \frac{1}{2}\varphi(v; j) + \frac{1}{2}\varphi(v_S^*; i) \Leftrightarrow \mu_\varphi(\frac{1}{2}(v; i) \oplus \frac{1}{2}(v_S^*; j)) = \mu_\varphi(\frac{1}{2}(v; j) \oplus \frac{1}{2}(v_S^*; i))$$

$$\Leftrightarrow \frac{1}{2}(v; i) \oplus \frac{1}{2}(v_S^*; j) \sim_\varphi (\frac{1}{2}(v; j) \oplus \frac{1}{2}(v_S^*; i).$$

(T/T*): Let $v, w \in SG_n$, $S \in M(v) \cap M(w)$ $(S \neq N)$, and $i \in N$

$$\Psi_i(v) - \Psi_i(v_S^*) = \Psi_i(w) - \Psi_i(w_S^*) \Leftrightarrow \Psi_i(v) + \Psi_i(w_S^*) = \Psi_i(w) + \Psi_i(v_S^*)$$

$$\Leftrightarrow \frac{1}{2}\varphi(v; i) + \frac{1}{2}\varphi(w_S^*; i) = \frac{1}{2}\varphi(w; i) + \frac{1}{2}\varphi(v_S^*; i)$$

$$\Leftrightarrow \mu_\varphi(\frac{1}{2}(v; i) \oplus \frac{1}{2}(w_S^*; i)) = \mu_\varphi(\frac{1}{2}(w; i) \oplus \frac{1}{2}(v_S^*; i)) \Leftrightarrow \frac{1}{2}(v; i) \oplus \frac{1}{2}(w_S^*; i) \sim_\varphi \frac{1}{2}(v_S^*; i) \oplus \frac{1}{2}(w; i).$$

(TGLB/AIVI) Let $v \in SG_n$, and $S \in M(v)$ $(S \neq N)$,

$$\sum_{i\in N} \Phi_i(v) = \sum_{i\in N} \Phi_i(v_S^*) \Leftrightarrow \sum_{i\in N} \varphi(v;i) = \sum_{i\in N} \varphi(v_S^*;i) \Leftrightarrow \sum_{i\in N} \frac{1}{n}\varphi(v;i) = \sum_{i\in N} \frac{1}{n}\varphi(v_S^*;i).$$

$$\Leftrightarrow \preceq(\bigoplus_{i\in N} \frac{1}{n}(v;i)) = \preceq(\bigoplus_{i\in N} \frac{1}{n}(v_S^*;i)) \Leftrightarrow \bigoplus_{i\in N} \frac{1}{n}(v;i) \sim_{\preceq} \bigoplus_{i\in N} \frac{1}{n}(v_S^*;i).$$

(AGLB/CIVI) Let $v \in SG_n$, and $S \in M(v)$ $(S \neq N)$,

$$\frac{1}{s}\sum_{i\in S}(\Phi_i(v) - \Phi_i(v_S^*)) = \frac{1}{n-s}\sum_{j\in N\setminus S}(\Phi_j(v_S^*) - \Phi_j(v))$$

$$\Leftrightarrow \frac{1}{s}\sum_{i\in S}(\varphi(v;i) - \varphi(v_S^*;i)) = \frac{1}{n-s}\sum_{j\in N\setminus S}(\varphi(v_S^*;j) - \varphi(v;j))$$

$$\Leftrightarrow \frac{1}{s}\sum_{i\in S}\varphi(v;i) + \frac{1}{n-s}\sum_{j\in N\setminus S}\varphi(v;j) = \frac{1}{s}\sum_{i\in S}\varphi(v_S^*;i) + \frac{1}{n-s}\sum_{j\in N\setminus S}\varphi(v_S^*;j)$$

$$\Leftrightarrow \preceq(\frac{1}{2}(\bigoplus_{i\in S}\frac{1}{s}(v;i)) \oplus \frac{1}{2}(\bigoplus_{i\in N\setminus S}\frac{1}{n-s}(v;i))) = \preceq(\frac{1}{2}(\bigoplus_{i\in S}\frac{1}{s}(v_S^*;i)) \oplus \frac{1}{2}(\bigoplus_{i\in N\setminus S}\frac{1}{n-s}(v_S^*;i)))$$

$$\Leftrightarrow \frac{1}{2}(\bigoplus_{i\in S}\frac{1}{s}(v;i)) \oplus \frac{1}{2}(\bigoplus_{i\in N\setminus S}\frac{1}{n-s}(v;i)) \sim_{\preceq} \frac{1}{2}(\bigoplus_{i\in S}\frac{1}{s}(v_S^*;i)) \oplus \frac{1}{2}(\bigoplus_{i\in N\setminus S}\frac{1}{n-s}(v_S^*;i)). \blacksquare$$

Now we can state the main result in this paper.

Theorem 3 There exists a unique von Neumann-Morgenstern ordering in $L(SG_n \times N)$ that satisfies An*, NP*, SymGL* (or T* or CED) and absolute indifference under the veil of ignorance (AIVI) (resp., conditional indifference under the veil of ignorance (CIVI)). Moreover it is the one represented by the utility function

$$sh(v;i) := Sh_i(v) \quad (\text{resp., } bz(v;i) := Bz_i(v)).$$

Proof. It is now an easy corollary of Theorem 2 and Propositions 1, 2 and 3. Let $\preceq$ be a preference ordering in $L(SG_n \times N)$ represented by $\varphi : SG_n \times N \to R$. And let Φ given by $\Phi_i(v) = \varphi(v;i)$, for all v in SG_n and all $i = 1,2,...,n$. Then, in view of Proposition 3, $\preceq$ satisfies An*, NP*, SymGL* (T*) and AIVI (resp., CIVI) if and only if Φ satisfies An, NP, SymGL (T) and TGLB (resp., AGLB). And by Theorem 2 this will be so if and only if it is $\Phi = \alpha Sh + \kappa 1$ (resp., $\Phi = \alpha Bz + \kappa 1$,) for some $\alpha > 0$ and $\kappa \in R$. In other words, if and only if $\varphi(v;i) = \alpha sh(v;i) + \kappa$ (resp., $\varphi(v;i) = \alpha bz(v;i) + \kappa$). Thus, if and only if $\preceq$ coincides with the VNM preferences represented by sh (resp., bz).

In case of assuming CED instead of T* or SymGL* note that, on the one hand, this condition implies T* (Proposition 1). Conversely, if $\preceq$ is the VNM preference on

$L(SG_n \times N)$ represented by $sh(v;i) := Sh_i(v)$ (the case $bz(v;i) := Bz_i(v)$ is entirely similar), let $l, l' \in L(SG_n \times N)$; with support in $SG_n \times \{i\}$ for some i, such that

$$\sum_{v \in SG_n} l(v;i)v = \sum_{v \in SG_n} l'(v;i)v.$$

The games on both sides of the equality are in $Co(SG_n)$, where Sh_i is defined and linear. Thus

$$\sum_{v \in SG_n} l(v;i)Sh_i(v) = \sum_{v \in SG_n} l'(v;i)Sh_i(v).$$

Or equivalently $\sum_{v \in SG_n} l(v;i)sh(v;i) = \sum_{v \in SG_n} l'(v;i)sh(v;i)$. That is to say $\overline{sh}(l) = \overline{sh}(l')$; i.e., $l \sim l'$. Thus μ verifies CED. ■

5 CONCLUDING REMARKS

As a first conclusion of this work we want to emphasize and vindicate Roth's (1977b, 1988) approach. We think that Roth's setting is particularly appropriate for a better understanding of the nature of power indices: Power indices should be understood as comparative assessments of the power attached to roles or positions in collective decision procedures modeled as simple superadditive games. In this context even the words "$\varphi_i(v)$ represents the power of player i in game v", though correct, can be misleading. What $\varphi_i(v)$ evaluates is the a priori capacity to influencing the outcome in any collective decision process modeled by game v of the player, whoever she or he be, that sits on seat i, or plays role i; in game v, and whatever the issue at stake. What are evaluated are roles or positions in collective decision procedures. To formalize this idea, Roth's setting seems the most suitable.

Roth's approach provides a point of view under which the non game-theoretical nature of power indices is conspicuous. There are no players, there are no strategies, no cake to share, no cooperation, no competition. Just roles in voting rules to be ranked. Even the TU framework is conceptually superfluous. When simple superadditive games are used to represent decision-making procedures, the only relevant information attached to the worth of a coalition is whether it is winning or not. The assumption of transferable utility that seems to underlie this model is unnecessary and out of place in this context. Consequently, if simple superadditive games are interpreted as models of voting rules, power indices should be interpreted and axiomatized accordingly, without attaching special meaning to the numerical worth of the coalitions beyond the dichotomy winning/losing. While solutions of cooperative TU games are aimed to give the utilities that players should expect, power measures are aimed to give a ranking of the roles or positions in collective decision procedures.

Apart from this basic coincidence with Roth's (1977b, 1988) approach we want to stress some discrepancies. As discussed in Section 3, Roth's interpretation of his di®erentiating axioms, "neutrality/aversion to strategic risk," is particularly inconsistent with the interpretation of power indices as utility functions representing expected utility preferences on roles in voting rules. On the other hand, the justi¯cation of "e±ciency" of the resulting utility function in this context for the Shapley-Shubik preference ordering remains unavoidably obscure in Roth's (1977b, 1988). It just appears as an unaccountable consequence of the other axioms.

Instead, the translation of our axioms, that characterize either index up to a zero and a unit in the domain of simple superadditive games in Laruelle and Valenciano (1999), into this setting turns out clear and natural. In particular, in our systems both di®erentiating axioms express two di®erent attitudes toward risk facing the veil of ignorance, in a context in which Rawls' concept seems especially suitable. Our "absolute indi®erence under the veil of ignorance" appears as the natural counterpart in Roth's setting of the up to now obscure "e±ciency" in this context, once stripped of its normalizing ingredients. While our "conditional indi®erence under the veil of ignorance" captures in similar terms the risk posture underlying the Banzhaf index.

It is not clear which of the two attitudes under the veil of ignorance is more plausible. But a direct consequence of absolute indi®erence under the veil of ignorance is that all games are indi®erent when all roles are equally probable. This entails that all symmetric games, in which all roles are interchangeable, are considered indi®erent. In other words, assuming anonymity*, absolute indi®erence under the veil of ignorance, the axiom that di®erentiates Shapley-Shubik's preferences, entails, for instance, that the all-player's unanimity decision rule and the simple majority rule are considered indi®erent. But this contradicts the usual common sense view.

Finally we want to point out some lines of further research. First, possibly the results presented here can be extended to semivalues and probabilistic values in general, in the domain of simple superadditive games at least. Second, usually simple games have received less attention than general TU games. Values and axiomatizations are ¯rst thought for the wider domain and only afterwards restricted or adapted to the particular case of simple games. It could be interesting for a change to do the opposite. Can our axioms be extended or adapted for general TU games in either the usual or Roth's set up? The answer is yes, but do they keep their characterizing power?

References

[1] Banzhaf, J., 1965, Weighted voting doesn't work : A Mathematical Analysis, Rutgers Law Review 19, 317-343.

[2] Blair, D. H., and R. P. McLean, 1990, Subjective Evaluations of n-Person Games, Journal of Economic Theory 50, 346-361.

[3] Dubey, P., 1975, On the Uniqueness of the Shapley Value, International Journal of Game Theory 4, 131-139.

[4] Dubey, P., and L. S. Shapley, 1979, Mathematical Properties of the Banzhaf Power Index, Mathematics of Operations Research 4, 99-131.

[5] Einy, E., 1987, Semivalues of Simple Games, Mathematics of Operations Research 12, 185-192.

[6] Hernstein, I. N., and J. Milnor, 1953, An Axiomatic Approach to Measurable Utility, Econometrica 21, 291-297.

[7] Laruelle, A., and F. Valenciano, 1998, On the measurement of inequality in the distribution of power in decision-making processes, working paper DT5/1998, Dpto. Econom¶a Aplicada I, Universidad del Pa¶s Vasco, Bilbao, Spain.

[8] Laruelle, A., and F. Valenciano, 1999, Shapley-Shubik and Banzhaf indices revisited. Discussion paper, No 2, Dpto. Econom¶a Aplicada IV, Universidad del Pa¶s Vasco, Bilbao, Spain.

[9] Rawls, A., 1972, A Theory of Justice, Oxford: Oxford University Press.

[10] Roth, A., 1977a, The Shapley Value as a von Neumann-Morgenstern Utility, Econometrica 45, 657-664.

[11] Roth, A., 1977b, Utility Functions for Simple Games, Journal of Economic Theory 16, 481-489.

[12] Roth, A., 1988, The Expected Utility of Playing a Game, in The Shapley Value. Essays in Honor of Lloyd S. Shapley, Edited by A. Roth, 51-70. Cambridge: Cambridge University Press.

[13] Shapley, L. S., and M. Shubik, 1954, A method for Evaluating the Distribution of Power in a Committee System, American Political Science Review 48, 787-792.

[14] Stra±n, P. D., 1977, Homogeneity, Independence and Power Indices, Public Choice 30, 107-118.

[15] von Neumann, J. and O. Morgenstern, 1944, 1947, 1953, Theory of Games and Economic Behavior. Princeton: Princeton University Press.

[16] Weber, R. J., 1988, Probabilistic Values for Games, in The Shapley Value. Essays in Honor of Lloyd S. Shapley, Edited by A. Roth, 101-119. Cambridge: Cambridge University Press.