Reprinted from JOURNAL OF ECONOMIC THEORY
All Rights Reserved by Academic Press, New York and London

Vol. 16, No. 2, December 1977
Printed in Belgium

Utility Functions for Simple Games*

This note characterizes the Shapley–Shubik index and the Banzhaf index on the class of simple games as utility functions which reflect different postures toward risk.

1. INTRODUCTION

A *simple game* on a set N of players is one in which the characteristic function v takes on only the values 0 and 1, and in which $v(S) = 1$ implies $v(T) = 1$ for all $S \subset T \subset N$. Such games arise naturally as models of political or economic situations in which every coalition of players is either "winning" or "losing."

In this note we investigate indices which reflect the relative power of each position (or player) in a simple game. We show that both the Shapley–Shubik index and the Banzhaf index correspond to von Neumann–Morgenstern utility functions, which differ only in their postures toward risk. The utility of playing a game from a given position is a reflection of the power which can be exerted by a player in that position. Chiefly, we apply the techniques developed in [4, 5] to the results presented in [8, 9].

A *game* on a finite universal set N of *positions* may be considered to be any function $v : 2^N \to R$ such that $v(\Phi) = 0$. In [10], the *value* of a game v is defined to be a vector-valued function $\phi(v) = (\phi_1(v),\ldots, \phi_n(v))$ which associates a real number $\phi_i(v)$ with each position $i \in N$, and which obeys the following conditions.

(1.1) For each permutation[1] $\pi : N \to N$, $\phi_{\pi i}(\pi v) = \phi_i(v)$.

(1.2) For each carrier[2] T of v, $\sum_{i \in T} \phi_i(v) = v(T)$.

(1.3) For any games v and w, $\phi(v + w) = \phi(v) + \phi(w)$.

* The author wishes to acknowledge the comments of D. Brown and P. Dubey.

[1] A permutation $\pi : N \to N$ is one-to-one and onto. The game πv is defined by $\pi v(\pi S) = v(S)$.

[2] A *carrier* of a game v is any coalition $T \subset N$ such that for all $S \subset N$, $v(S) = v(S \cap T)$. The smallest carrier in a game may be viewed as the set of active players in the game.

ISSN 0022-0531

Shapley showed that the unique value defined on all games has the form

$$\phi_i(v) = \sum_{S \subset N} \left(\frac{(s-1)!\,(n-s)!}{n!}\right)(v(S) - v(S-i)),$$

where s and n denote the cardinality of the sets S and N. In [8] it was shown that this value can be viewed as a utility function which expresses the desirability of playing a given position in a game.

In [11] the Shapley value is studied in the context of simple games. Observe that if v is a simple game, then the quantity $(v(S) - v(S-i))$ equals 0 unless S is a winning coalition and $(S - i)$ is a losing coalition, in which case it equals 1. Consequently, if we suppose that players in a simple game v "vote" in random order, then $\phi_i(v)$ is precisely the probability that player i will cast a "pivotal" vote. As such, it can be viewed as an apriori index of power in simple games, and is referred to as the Shapley–Shubik index.

However, if only simple games are to be considered, conditions (1.1), (1.2), and (1.3) no longer specify a *unique* functional form. This is because condition (1.3) becomes nonbinding since the class of simple games is not closed under addition. (So if v and w are nontrivial simple games, $v(N) = w(N) = 1$, and the game $v + w$ is not simple, since $v(N) + w(N) = 2$.)

Another value for simple games which has received attention in the literature is the Banzhaf index, first introduced in [1]. The Banzhaf index takes as a measure of power the relative ability of players to transform winning coalitions into losing coalitions, and vice versa. Define a *swing* for position $i \in N$ as a pair $(S, S - i)$ such that the coalition S is winning, and $S - i$ is losing (i.e., $v(S) = 1$ and $v(S - i) = 0$). Let $\eta_i(v)$ denote the number of swings for position i in game v, and let $T(v) = \sum_{i \in N} \eta_i(v)$. Then the Banzhaf index of relative power for each position is

$$\beta_i(v) = \eta_i(v)/T(v) \qquad \text{for } i = 1,\ldots, n.$$

We refer to $\eta_i(v)$ as the nonnormalized Banzhaf index.

The Shapley–Shubik and the (normalized or nonnormalized) Banzhaf indices yield different rankings of the positions in a given simple game. Consequently, it is desirable to find a common interpretation of the indices which will permit us to investigate their differences and similarities. This task is facilitated by the following two propositions which are presented in [4, 5].

PROPOSITION 1 (Dubey). *The Shapley–Shubik index is the unique function ϕ defined on simple games which satisfies conditions* (1.1) *and* (1.2) *and which has the property that*:

(1.4) *For any simple games v, w, $\phi(v \vee w) + \phi(v \wedge w) = \phi(v) + \phi(w)$ where the games $(v \vee w)$ and $(v \wedge w)$ are defined by*

$$(v \vee w)(S) = 1 \quad \textit{if } v(S) = 1 \textit{ or } w(S) = 1,$$
$$= 0 \quad \textit{otherwise},$$

and

$$(v \wedge w)(S) = 1 \quad \textit{if } v(S) = 1 \textit{ and } w(S) = 1,$$
$$= 0 \quad \textit{otherwise}.$$

PROPOSITION 2 (Dubey). *The nonnormalized Banzhaf index is the unique function defined on simple games which satisfies the following four conditions.*

(1.5) *If* $i \in N$ *is a dummy*[3] *in* v, *then* $\eta_i(v) = 0$.

(1.6) $\sum_{i \in N} \eta_i(v) = T(v)$.

(1.7) *For each permutation* $\pi : N \to N$, $\eta_{\pi i}(\pi v) = \eta_i(v)$.

(1.8) *For any simple games,* v, w, $\phi(v \vee w) + \phi(v \wedge w) = \phi(v) + \phi(w)$.

The introduction of condition (1.4) by Dubey permits these indices to be studied in the context of simple games alone. This is useful here since it permits us to express the utility for playing a simple game in terms of preferences which involve only simple games.

In the next section, we use Propositions 1 and 2 to show that ϕ, η, and β can be viewed as cardinal utility functions which differ only in their postures toward risk. It will be seen that conditions (1.2) or (1.6) express a posture toward one kind of risk, while the somewhat opaque conditions (1.4) (or (1.8)) express a posture toward another kind of risk.

2. UTILITY FUNCTIONS FOR SIMPLE GAMES

Let C be the class of simple games defined on a finite set N, and let M be the mixture space generated by $C \times N$. Then the elements of M are elements (w, i) of $C \times N$, and lotteries of the form $[p(w, i); (1 - p)(v, j)]$ where (w, i) and (v, j) are in $C \times N$, and p is a probability (i.e., $p \in [0, 1]$).[4] An individual involved in such a lottery will, with probability p, play position i in game w, and with probability $(1 - p)$ play position j in game v.

[3] A *dummy* in a game v is an $i \in N$ for which there are no swings.

[4] We assume the usual properties of mixture spaces, i.e., for all elements $a, b \in M$ and all probabilities p and q, we have

$$[1a; (1 - 1)b] = a;$$
$$[pa; (1 - p)b] = [(1 - p)b; pa];$$

and

$$[q[pa; (1 - p)b]; (1 - q)b] = [qpa; (1 - qp)b].$$

Let P be a (strict) preference relation defined on M.[5] (Read $(w, i)\ P(v, j)$ as "it is preferable to play position i in game w than to play position j in game v"). We take P to be *continuous* on M; i.e., if $a, b, c \in M$ such that $aPbPc$, then there exists a (unique) $q \in (0, 1)$ such that $bI[qa; (1 - q)\ c]$.

Denote by v_R and v_0 the games defined by

$$v_R(S) = 1 \quad \text{if } R \subset S; \qquad v_0(S) = 0 \text{ for all } S \subset N,$$
$$= 0 \quad \text{otherwise.}$$

For each $i \in N$ denote by $D_i \subset C$ the set of simple games for which player i is a dummy. We take P to have the following properties.

(2.1) For all $v \in C$, $i \in N$ and every permutation $\pi : N \to N$, $(v, i)\ I(\pi v, \pi i)$.

(2.2) For every $i \in N$, $v \in D_i$ implies $(v, i)\ I(v_0, i)$ and $(v_i, i)\ P(v_0, i)$. For every $v \in C$, $(v_i, i)\ R(v, i)\ R(v_0, i)$.

It is well known (cf. [6]) that such a preference can be represented by a cardinal utility function θ; i.e., a function θ such that for all $a, b \in M$

$$\theta(a) > \theta(b) \qquad \text{iff } aPb,$$

and

$$\theta([pa; (1 - p)\ b]) = p\theta(a) + (1 - p)\ \theta(b).$$

Furthermore, θ is unique up to an affine transformation, so we can set $\theta(v_i, i) = 1$ and $\theta(v_0, i) = 0$. For an arbitrary element (v, i) of $C \times N$ we have

$$\theta_i(v) \equiv \theta(v, i) = q,$$

where q is the number such that

$$(v, i)\ I[q(v_i, i); (1 - q)(v_0, i)].$$

By the continuity of P and condition (2.2), $\theta_i(v)$ is well defined.

We have yet to completely specify the preference P. We do so by expressing the preferences involving two kinds of risk.

[5] For $a, b \in M$ we write aIb (a is indifferent to b) if neither aPb nor bPa. We write aRb if either aPb or aIb, and assume that R is a transitive, complete order on M. Furthermore, we assume that if aIb, then for every $c \in M$ and $p \in [0, 1]$,

$$[pa; (1 - p)c]I[pb; (1 - p)c].$$

(2.3) Ordinary risk neutrality: For all simple games v, w

$$[\tfrac{1}{2}(v, i); \tfrac{1}{2}(w, i)]\, I [\tfrac{1}{2}((v \vee w), i); \tfrac{1}{2}((v \wedge w), i)].$$

(2.4) Strategic risk neutrality: For all $R \subset N$ and $i \in R$,

$$(v_R, i)\, I \left[\frac{1}{r}(v_i, i); \left(1 - \frac{1}{r}\right)(v_0, i)\right].$$

Condition (2.3) specifies indifference between two lotteries. One lottery results in either the game v or the game w, while the other results in either the game $(v \vee w)$ or the game $(v \wedge w)$. Note that the condition is plausible, since $(v \vee w) \geqslant v$ and $(v \wedge w) \leqslant w$. Thus, it is to be expected that it is more desirable to play in the game $(v \vee w)$ than to play the same position in game v, and less desirable to play in $(v \wedge w)$ than in w. Condition (2.3) expresses the intensity of these desirablility comparisons. Note also that a given coalition S has the same probability of being a winning coalition in either lottery.[6]

Condition (2.4) specifies indifference between playing the game v_R as one of r players in the unique minimal winning coalition, or participating in a lottery which gives probability $1/r$ of being a dictator and probability $[1 - (1/r)]$ of being a dummy. Note that the risk involved in playing the game v_R is strategic rather than probabilistic—no gamble is involved.

We can now state the following theorem.

THEOREM 1. *If P is a preference obeying conditions* (2.1) *through* (2.4), *then the unique utility θ such that $\theta_i(v_i) = 1$ and $\theta_i(v_0) = 0$ is equal to the Shapley–Shubik index.*

LEMMA 1. *If P obeys condition* (2.1), *then, for every $v \in C$, $i \in N$ and permutation $\pi : N \to N$,*

$$\theta_i(v) = \theta_{\pi i}(\pi v).$$

Proof. Immediate from (2.1) and the definition of utility.

LEMMA 2. *If P obeys conditions* (2.1), (2.2), *and* (2.4), *then for each $R \subset N$,*

$$\begin{aligned} \theta_i(v_R) &= \frac{1}{r} \qquad \textit{if } i \in R, \\ &= 0 \qquad \textit{if } i \notin R. \end{aligned}$$

Proof. If $i \notin R$ then $(v_R, i)\, I (v_0, i)$ by (2.2), and so $\theta_i(v_R) = \theta_i(v_0) = 0$. If $i \in R$, then $\theta_i(v_R) = 1/r$ by (2.4) and the definition of the utility θ.

[6] This latter observation was pointed out to me by Lloyd Shapley.

LEMMA 3. *If P obeys condition* (2.3), *then*

$$\theta_i(v \vee w) + \theta_i(v \wedge w) = \theta_i(v) + \theta_i(w).$$

Proof. From the definition of utility,

$$\theta[\tfrac{1}{2}((x \vee w), i); \tfrac{1}{2}((v \wedge w), i)] = \tfrac{1}{2}\theta((v \vee w), i) + \tfrac{1}{2}\theta(v \wedge w), i)$$

and

$$\theta[\tfrac{1}{2}(v, i); \tfrac{1}{2}(w, i)] = \tfrac{1}{2}\theta(v, i) + \tfrac{1}{2}\theta(w, i).$$

Consequently, by condition (2.3), we have

$$\tfrac{1}{2}\theta(v, i) + \tfrac{1}{2}\theta(w, i) = \tfrac{1}{2}\theta((v, \vee w), i) + \tfrac{1}{2}\theta((v \wedge w), i).$$

So far we have demonstrated that θ obeys conditions (1.1) and (1.4), and that for every $R \subset N$, $\theta(v_R) = \phi(v_R)$; i.e., θ coincides with the Shapley–Shubik index on the games v_R. (Note that conditions (1.1) and (1.2) determine the value of $\phi(v_R)$.) To complete the proof of the theorem, we show that θ coincides with ϕ on every game $v \in C$.

Proof of Theorem 1. Let $v \in C$, and let $R_1, R_2, \ldots, R_k \subset N$ be all the distinct minimal winning coalitions[7] of v. Then we say the game v is in *class k*, and note that

$$v = v_{R_1} \vee v_{R_2} \vee \cdots \vee v_{R_k}.$$

If v is in class $k = 0$, then $v = v_0$ and $\theta(v) = \phi(v) = 0$. If v is in class $k = 1$, then $v = v_R$, and $\theta(v)$ is defined by Lemma 2 and is equal to $\phi(v)$.

Suppose that for games v in classes $k = 1, 2, \ldots, m$ it has been shown that θ is well defined and coincides with the Shapley–Shubik index. Consider a game v in class $m + 1$. Then

$$v = v_{R_1} \vee v_{R_2} \vee \cdots \vee v_{R_m} \vee v_R = w \vee v_R,$$

where w is a game in class m.

So, by Lemma 3,

$$\theta_i(v) = \theta_i(w \vee v_R) = \theta_i(w) + \theta_i(v_R) - \theta_i(w \wedge v_R).$$

But we show that the game $(w \wedge v_R)$ cannot be in a higher class than w, so by the inductive hypothesis the terms on the right-hand side of the above

[7] A coalition $R \subset N$ is *minimal winning* in v if $v(R) = 1$ and if $S \subset R$, $S \neq R$ implies $v(S) = 0$.

expression are uniquely determined and equal to the Shapley–Shubik index. Consequently, (from property (1.4) of ϕ) we will have shown that $\theta(v) = \phi(v)$ for all simple games v.

To see that the game $w' = (w \wedge v_R)$ cannot be in a higher class than the game w, consider a minimal winning coalition S' of the game w'. By the definition of w' we know that $S' \supset R$ and $w(S') = 1$. If $S' = R$, then $w' = v_R$ and we are done (since except for the game v_0, every game has at least one minimal winning coalition). Otherwise, $S' = S \cup R$, where S is a minimal winning coalition in the game w. (Of course, S and R need not be disjoint.)

Consider now a coalition T' which is minimal winning in w'. Then $T' = T \cup R$ where T is minimal winning in w. If $T' \neq S'$, then $T \neq S$. Consequently, every minimal winning coalition in w' can be identified with a *distinct* minimal winning coalition in w, so w' cannot be in a higher class than w. This completes the proof.[8]

So the Shapley–Shubik index is the utility function representing preferences described by conditions (2.1) through (2.4). Naturally, different preferences will give rise to different functions. Suppose, for instance, that the posture toward strategic risk is represented not by condition (2.4) but by the following condition for every $R \subset N$ and $i \in R$.

$$\left[\frac{1}{T(v_R)}(v_R, i); \left(1 - \frac{1}{T)v_R)}\right)(v_0, i)\right]$$

$$I\left[\frac{1}{r2^{n-1}}(v_i, i); \left(1 - \frac{1}{r2^{n-1}}\right)(v_0, i)\right]. \tag{2.5}$$

Then the following theorem says that the nonnormalized Banzhaf index is a cardinal utility for the preference relation P.

THEOREM 2. *If P is a preference obeying conditions* (2.1), (2.2), (2.3), *and* (2.5), *then the unique utility θ such that $\theta_i(v_i) = 2^{n-1}$ and $\theta_i(v_0) = 0$ is equal to the normalized Banzhaf index.*

The proof is precisely like the proof of Theorem 1, once it has been observed that condition (2.5) implies that

$$\begin{aligned} \theta_i(v_R) = \eta_i(v_R) &= \frac{T(v_R)}{r} && \text{for } i \in R, \\ &= 0 && \text{for } i \notin R. \end{aligned}$$

So the nonnormalized Banzhaf index and the Shapley–Shubik index reflect preferences which differ only in their postures toward strategic risk.

[8] The proof presented here, which requires only one induction step, resembles the proof used in [5] to prove Proposition 1. That proof required two induction steps.

Similarly, it is not difficult to show that the ordinary (normalized) Banzhaf index corresponds to preferences which obey condition (2.4) but not condition (2.3). That is, the Banzhaf index reflects preferences which are neutral to strategic risk, but not to ordinary risk. The normalization has the effect of changing the risk posture, since each game is normalized independently (i.e. each game v is normalized by $T(v)$.)

3. Discussion

We have seen that the difference between the Shapley–Shubik index and the nonnormalized Banzhaf index results from different postures toward strategic risk. That is, the two indices reflect different attitudes toward the relative benefits of engaging in strategic interaction with other players in games of the form v_R.

The difference between the Shapley–Shubik index and the ordinary Banzhaf index, on the other hand, reflects different postures towards ordinary risk—the kind which results from lotteries, rather than from strategic interactions. Thus, the difference between these two indices seems to be essentially non-game-theoretic (cf. [3, pp. 195–196]).

It is natural to consider a whole spectrum of risk postures, and to examine the resulting utility functions. It seems likely that this point of view will serve to illuminate some of the recent work (cf. [2, 7]) on alternative formulations of the value concept by permitting the axioms involved to be interpreted as resulting from choices based on well-defined preferences which explicitly reveal the assumptions involving risk,

References

1. J. F. Banzhaf, III, Weighted voting doesn't work: A mathematical analysis, *Rutgers Law Rev.* **19** (1965), 317–343.
2. D. H. Blair, Essays in social choice theory, Ph.D. thesis, Yale University, 1976.
3. S. Brams, "Game Theory and Politics," Free Press, New York, 1975.
4. P. Dubey, Some results on values of finite and infinite games, Ph.D. thesis, Cornell University, 1975.
5. P. Dubey, On the uniqueness of the Shapley value, *Intern. J. Game Theory* **4**, No. 3 (1975), 131–139.
6. I. N. Herstein and J. Milnor, An axiomatic approach to measurable utility, *Econometrica* **21** (1953), 291–297.
7. R. Meyerson, A theory of cooperative games, Ph.D. dissertation, Harvard University, 1976.
8. A. E. Roth, The Shapley value as a von Neumann–Morgenstern utility, *Econometrica* **45** (1977), 657–664.

9. A. E. ROTH, Bargaining ability, the utility of playing a game, and models of coalition formation, *J. Math. Psych.*, in press.
10. L. S. SHAPLEY, A value for *n*-person games, *Ann. Math. Studies* **28** (1953), 307–318.
11. L. S. SHAPLEY AND M. SHUBIK, A method for evaluating the distribution of power in a committee system, *Amer. Polit. Rev.* **48** (1954), 787–792.

RECEIVED: March 16, 1976; REVISED: May 3, 1977

ALVIN E. ROTH

Departments of Business Administration and Economics
University of Illinois
Urbana, Illinois 61801

Printed by the St Catherine Press Ltd., Tempelhof 37, Bruges, Belgium.