Characterization of the Banzhaf Value
Using a Consistency Axiom

Joss Erick Sánchez Pérez
Facultad de Economía, UASLP,
Av. Pintores s/n, Col. B. del Estado 78213,
San Luis Potosí, México
email: joss.sanchez@uaslp.mx

ABSTRACT

One of the important properties characterizing cooperative game solutions is consistency. This notion establishes connections between the solution vectors of a cooperative game and those of its reduced game. The last one is obtained from the initial game by removing one or more players and by giving them payoffs according to a specific principle. Consistency of a solution means that the restriction of a solution payoff vector of the initial game to any coalition belongs to the solution set of the corresponding reduced game. In this paper, we show that the Banzhaf value is consistent with respect to a suitable reduced game. We also show that consistency together with standardness for two-person games characterize the Banzhaf value.

RESUMEN

Una de las propiedades importantes caracterizando soluciones de juegos cooperativos es consistencia. Esta noción establece conexiones entre los vectores solución de un juego cooperativo y aquellos de su juego reducido. Este último es obtenido a partir del juego inicial al remover uno o más jugadores y asignándoles pagos de acuerdo a un principio específico. Consistencia de una solución significa que la restricción de una solución del juego inicial a cualquier coalición, pertenece al conjunto de soluciones del correspondiente juego reducido. En este artículo, mostramos que consistencia junto con estandarización para juegos de dos personas caracterizan el valor de Banzhaf.

1   Introduction

The mathematical approach to a proposed solution is to examine a number of its (elementary) properties and, if possible, to provide a minimal number of properties which fully characterize the solution. In their well known paper, Hart and Mas-Colell (1989) characterized the Shapley value for the class of TU cooperative games by means of the following axioms[1]:

i) Consistency.

ii) Standardness for two-person games.

A cooperative game is always described by a finite player set as well as a real-valued "characteristic function" on the collection of subsets of the player set. A so-called reduced game is deducible from a given cooperative game by removing one or more players on the understanding that the removed players will be paid according to a specific principle (e.g., a proposed payoff vector). The remaining players form the player set of the reduced game: the characteristic function of which is composed of the original characteristic function, the proposed payoff vector, and / or the solution in question. The consistency property for the solution states that if all the players are supposed to be paid according to a payoff vector in the solution set of the original game, then the players of the reduced game can achieve the correspondig payoff vector in the solution set of the reduced game. In other words, there is no inconsistency in what the players of the reduced game can achieve, in either the original game or the reduced game. This paper presents, using a standard notion of consistency, that the Banzhaf value is the unique consistent extension which "divides the surplus equally" in two-person games.

2   Preliminaries

By an *n-person game* in characteristic function form (or a TU cooperative game), in what follows just a game, we mean a pair (N, v), where $N = \{1, ..., n\}$ is a finite set of players and v is a function $v : 2^N \to \mathbb{R}$ with the property that $v(\phi) = 0$ (2^N denotes the set of subsets of N). We usually refer to subsets S of N as *coalitions* and to the number $v(S)$ as the *worth* of S. A game (N, v) is superadditive if $v(S \cup T) \geq v(S) + v(T)$ for all $S, T \subset N$, and it is subadditive if the inequality holds in the other direction. There are several interpretations for (N, v), it depends on what people want to model. For instance, if the game is superadditive, $v(S)$ means the maximal amount the players

---

[1]The precise definitions will be provided in Section 3.

in S can get if they decide to play together. While the game is subadditive, $v(S)$ usually means the joint cost that players in S have to pay to get a service if they hired the service together. Additionally, we will denote the cardinality of a set by its corresponding lower-case letter, for instance $n = |N|$, $s = |S|$, $t = |T|$, and so on.

We denote by G^N the set of all games with a fixed set of players N, i.e., $G^N = \{v : 2^N \to \mathbb{R} \mid v(\phi) = 0\}$. A solution $\varphi : G^N \to \mathbb{R}^n$ in G^N is a rule to divide the common gain or cost among the players in N. Let Γ^N be the set of solutions in G^N.

Given $v, w \in G^N$ and $\lambda \in \mathbb{R}$, we define the sum and the product $v + w$ and λv in G^N in the usual form, i.e.,

$$(v + w)(S) = v(S) + w(S) \quad \text{and} \quad (\lambda v)(S) = \lambda v(S)$$

respectively.

In a similar manner, for $\varphi, \psi \in \Gamma^N$ and $\lambda \in \mathbb{R}$, we define the sum and the product $\varphi + \psi$ and $\lambda \varphi$ in G^N by

$$(\varphi + \psi)(v + w) = \varphi(v) + \psi(w) \quad \text{and} \quad (\lambda \varphi)(v) = \lambda \varphi(v)$$

It is easy to verify that G^N and Γ^N are vector spaces with these operations.

Next, we define some axioms which are common in game theory.

Axiom 1 (Linearity). *The solution φ is linear if $\varphi(v + w) = \varphi(v) + \varphi(w)$ and $\varphi(cv) = c\varphi(v)$, for all $v, w \in G^N$ and $c \in \mathbb{R}$.*

Let us consider the group $S_n = \{\theta : N \to N \mid \theta \text{ is bijective}\}$, the group of permutations of the set of players. For every $\theta \in S_n$ and $v \in G^N$ we define another game $(N, \theta \cdot v)$ as follows

$$\theta \cdot v(S) = v(\theta^{-1}S)$$

Axiom 2 (Symmetry). *The solution φ is said to be symmetric if and only if $\varphi(\theta \cdot v) = \theta \cdot \varphi(v)$ for every $\theta \in S_n$ and $v \in G^N$.*

The player i is said to be a dummy player in the game (N, v) if $v(S \cup \{i\}) = v(S)$ for every $S \subseteq N$. The dummy axiom requests that every dummy player in (N, v) gets a zero payoff.

Axiom 3 (Dummy). *If the player i is a dummy player in (N, v) then $\varphi_i(v) = 0$.*

For every coalition $T \subseteq N$ we obtain a new game v_T where the set of players T is considered as one player; we denote this player by T. The space of players for v_T is $N \backslash T \cup \{T\}$ and is defined by:

$$v_T(S) = v(S) \quad \text{and} \quad v_T(S \cup \{T\}) = v(S \cup T)$$

where $S \subseteq N \backslash T$. The reduction axiom says that unification of any two players has to be profitable (or at least never harmful).

Axiom 4 (Reduction). *The solution φ satisfies the reduction axiom if and only if $\varphi_i(v) + \varphi_j(v) \leq \varphi_T(v_T)$ for every coalition $T = \{i, j\}$ of two players.*

Theorem 5 (Lehrer, 1988). *There exists a unique solution φ that satisfies linearity, symmetry, dummy and reduction axioms. Furthermore it is given by*

$$\varphi_i(v) = \frac{1}{2^{n-1}} \sum_{S \subseteq N \backslash \{i\}} [v(S \cup \{i\}) - v(S)]$$

3   Consistency and the Main Result

This section is devoted to a characterization of the Banzhaf value by means of an internal consistency property. Intuitively speaking, let φ be a function that associates a payoff to every player in every game and suppose that one player leaves a game with some payoff. Reduced games describe what games are played between the remaining players, i.e. what is earned by coalitions of the remaining players after one player has left the game. Then φ is said to be consistent if, when it is applied to any reduced game, it yields the same payoffs as in the original game.

Definition 6. *A value $\varphi : G^N \to \mathbb{R}^n$ is consistent with respect to the reduced game $(N \backslash \{i\}, v^\varphi_{N \backslash \{i\}})$ if, for $N \backslash \{i\} \subset N$ and all $v \in G^N$, one has*

$$\varphi_j(N \backslash \{i\}, v^\varphi_{N \backslash \{i\}}) = \varphi_j(N, v) \qquad for \ all \ j \in N \backslash \{i\}$$

The various consistency requirements differ in the precise definition of the reduced game (i.e., exactly how the players outside are being paid off). It is well known that different solutions are consistent with respect to different reduced games. Next, we will see that the Banzhaf value is consistent according to the following reduced game concept:

Definition 7. *Given a solution $\varphi : G^N \to \mathbb{R}^n$ in G^N, a game (N, v) (with $n \geq 3$). The reduced game $(N \backslash \{k\}, v^\varphi_{N \backslash \{k\}})$ is defined by*

$$v^\varphi_{N \backslash \{k\}}(S) = \begin{cases} \frac{1}{2}[v(N) + v(N \backslash \{k\}) - v(\{k\})] & if \ S = N \backslash \{k\} \\ \frac{1}{2}[v(S) - v(N \backslash S)] & if \ \phi \neq S \subset N \backslash \{k\} \\ 0 & if \ S = \phi \end{cases} \qquad (1)$$

The first result is as follows:

Proposition 8. *The Banzhaf value is a consistent solution function.*

Proof. We must apply this value to the reduced game $(N \backslash \{k\}, v^\varphi_{N \backslash \{k\}})$. We note that player j in $(N \backslash \{k\}, v^\varphi_{N \backslash \{k\}})$ gets:

$$\varphi_j(N \backslash \{k\}, v^\varphi_{N \backslash \{k\}}) = \frac{1}{2^{n-2}} \sum_{S \subseteq N \backslash \{j,k\}} \left[v^\varphi_{N \backslash \{k\}}(S \cup \{j\}) - v^\varphi_{N \backslash \{k\}}(S) \right]$$

Then, using the reduced game concept (1) in the last expression, we obtain:

$$
\begin{aligned}
2^{n-2} \cdot \varphi_j(N \backslash \{k\}, v^{\varphi}_{N \backslash \{k\}}) &= \sum_{S \subseteq N \backslash \{j,k\}} \left[v^{\varphi}_{N \backslash \{k\}}(S \cup \{j\}) - v^{\varphi}_{N \backslash \{k\}}(S) \right] \\
&= \sum_{\phi \neq S \subset N \backslash \{j,k\}} [v(S \cup \{j\}) - v(N \backslash (S \cup \{j\})) - v(S) + v(N \backslash S)] \\
&\quad + v(\{j\}) - v(N \backslash \{j\}) + v(N) + v(N \backslash \{k\}) - v(N \backslash \{j,k\}) + v(\{j,k\}) \\
&= \frac{1}{2} \sum_{S \subseteq N \backslash \{j,k\}} [v(S \cup \{j\}) - v(S) + v(N \backslash S) - v(N \backslash (S \cup \{j\}))]
\end{aligned}
$$

We observe that in the sum, the worth $v(N \backslash S)$ corresponds to the worths of coalitions containing $\{j,k\}$; while the worth $v(N \backslash S \cup \{j\})$ corresponds to the worths of coalitions containing $\{k\}$ but not $\{j\}$. Hence, we can apply the change:

$$
\begin{aligned}
\varphi_j(N \backslash \{k\}, v^{\varphi}_{N \backslash \{k\}}) &= \frac{1}{2^{n-1}} \left\{ \sum_{S \subseteq N \backslash \{j,k\}} [v(S \cup \{j\}) - v(S)] + \sum_{S \subseteq N \backslash \{j\}, S \ni \{k\}} [v(S \cup \{j\}) - v(S)] \right\} \\
&= \frac{1}{2^{n-1}} \sum_{S \subseteq N \backslash \{j\}} [v(S \cup \{j\}) - v(S)] \\
&= \varphi_j(N, v)
\end{aligned}
$$

$\square$

Next, we will show that the property of consistency almost characterizes the Banzhaf value. "Almost" refers to the "initial conditions", namely, the behavior of the solution for two-person games. The next theorem establishes the main result of this work, we show that consistency together with standardness for two-person games characterize the Banzhaf value. This characterization is equivalent to that of the Shapley value given by Hart and Mas-Colell (1989) where the only difference rests upon the reduced game concept underlying the notion of consistency. First, we recall that a solution φ is *standard for two-person games* if

$$
\varphi_i(\{i,j\}, v) = v(\{i\}) + \frac{1}{2} [v(\{i,j\}) - v(\{i\}) - v(\{j\})]
$$

for all $i \neq j$ and all v. Thus, the surplus $[v(\{i,j\}) - v(\{i\}) - v(\{j\})]$ is equally divided among the two players.

Theorem 9. *Let $\varphi : G^N \to \mathbb{R}^n$ in G^N be a solution function. Then:*

    (i) φ is consistent with respect to the reduced game(1); and

    (ii) φ is standard for two-person games;

    if and only if φ is the Banzhaf value.

Proof. One direction is immediate, recall Proposition 8 and the fact that:

$$
\begin{aligned}
\varphi_i(\{i,j\},v) &= \frac{1}{2}\sum_{S\subseteq\{j\}}[v(S\cup\{i\})-v(S)] \\
&= v(\{i\})+\frac{1}{2}[v(\{i,j\})-v(\{i\})-v(\{j\})]
\end{aligned}
$$

For the other direction, the proof is by induction on the number of players. Assume $\varphi,\psi: G^N \to \mathbb{R}^n$ satisfy (i) and (ii). We claim first that $\varphi=\psi$ for two-person games, since for $k\in\{i,j\}$:

$$
\begin{aligned}
\varphi_k(\{i,j\},v) &= v(\{k\})+\frac{1}{2}[v(\{i,j\})-v(\{i\})-v(\{j\})] \\
&= \frac{1}{2}\sum_{\substack{S\ni k \\ S\subseteq\{i,j\}}}[v(S)-v(\{i,j\}\backslash S)] \\
&= \psi_k(\{i,j\},v)
\end{aligned}
$$

Now, suppose that φ and ψ coincide for all the games with $n-1$ players. The two-person reduced games: $(\{i,j\},v^\varphi_{\{i,j\}})$ and $(\{i,j\},v^\psi_{\{i,j\}})$ coincide since, $v^\varphi_{\{i,j\}}(S)=v^\psi_{\{i,j\}}(S)$ for any $S\subset\{i,j\}$.

Then $\varphi_i(v^\varphi_{\{i,j\}})=\psi_i(v^\psi_{\{i,j\}})$ by standarness for two-person games, and $\varphi_i(N,v)=\varphi_i(\{i,j\},v^\varphi_{\{i,j\}})=\psi_i(\{i,j\},v^\psi_{\{i,j\}})=\psi_i(N,v)$ by consistency.

In a similar manner, $\varphi_j(N,v)=\psi_j(N,v)$. Applying this to any pair of players, one gets $\varphi_i(N,v)=\psi_i(N,v)\ \ \forall i\in N$. □

References

[1] HART, S. AND MAS-COLELL, A., *Potencial, Value, and Consistency*, Econometrica, **57** (1989), 589–614.

[2] LEHRER, E., *An axiomatization of the Banzhaf value*, International Journal of Game Theory, **17** (1988), 89–99.

[3] OWEN, G., *Consistency in values*, Sixth Spanish Meeting on Game Theory and Practice, (2003), 15–30.