UNIVERSIDAD CARLOS III DE MADRID

working papers

Working Paper 05-73
Economics Series 39
November 2005

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624 98 75

MARGINAL CONTRIBUTIONS AND EXTERNALITIES IN THE VALUE *

Geoffroy de Clippel [1] and Roberto Serrano [2]

Abstract

For games in partition function form, we explore the implications of distinguishing between the concepts of intrinsic marginal contributions and externalities. If one requires efficiency for the grand coalition, we provide several results concerning extensions of the Shapley value. Using the axioms of efficiency, anonymity, marginality and monotonicity, we provide upper and lower bounds to players' payoffs when affected by external effects, and a characterization of an ``externality-free" value. If the grand coalition does not form, we characterize a payoff configuration on the basis of the principle of balanced contributions. We also analyze a game of coalition formation that yields sharp predictions

Keywords:. externalities, marginal contributions, Shapley value, Pigouvian transfers, coalition formation

JEL classification:. C7, D62

* **Acknowledgements**: Serrano gratefully acknowledges Universidad Carlos III and CEMFI in Madrid for their hospitality, and Fundacion Banco Herrero and Universidad Carlos III for research support

[1] Rice University. Departament of Economics. clippel@rice.edu
[2] Brown University. Department of Economics, Providende, RI 02912, U.S.A. e-mail: roberto_serrano@brown.edu

Marginal Contributions and Externalities in the Value*

Geoffroy de Clippel[†] and Roberto Serrano[‡]

November 2005

Abstract

For games in partition function form, we explore the implications of distinguishing between the concepts of intrinsic marginal contributions and externalities. If one requires efficiency for the grand coalition, we provide several results concerning extensions of the Shapley value. Using the axioms of efficiency, anonymity, marginality and monotonicity, we provide upper and lower bounds to players' payoffs when affected by external effects, and a characterization of an "externality-free" value. If the grand coalition does not form, we characterize a payoff configuration on the basis of the principle of balanced contributions. We also analyze a game of coalition formation that yields sharp predictions.

JEL classification numbers: C7, D62
Keywords: externalities, marginal contributions, Shapley value, Pigouvian transfers, coalition formation.

1 Introduction

Since the path-breaking work of Shapley (1953), much effort has been devoted to the problem of finding a unique "fair" distribution of the surplus generated by a collection of people that are willing to cooperate with one another. More recently, the same question has been posed in the realistic case where externalities across coalitions are present. This is the general problem to which this paper contributes. In tackling the question, we find that sorting out the effects of *intrinsic marginal contributions* of players to coalitions from those coming from *externalities* is helpful to obtain meaningful answers.[1]

*Serrano gratefully acknowledges the hospitality of Universidad Carlos III and CEMFI in Madrid, and the research support from Fundación Banco Herrero and Universidad Carlos III.

[†]Department of Economics, Rice University; clippel@rice.edu

[‡]Department of Economics, Brown University; roberto_serrano@brown.edu

[1]William Thomson (private communication) argues that in the Shapley value literature we should avoid redundancies and speak of "contributions" instead of "marginal contributions." His contribution is not marginal. We will side with the majority for now, and ours will be a "marginal contribution."

Shapley (1953) studied games in which there are no externalities across coalitions. In such settings, Shapley characterized a unique solution using the axioms of efficiency, anonymity, additivity and null player. Today we refer to this solution as the Shapley value, which happens to be calculated as the average of marginal contributions of players to coalitions. This comes as a surprise at first glance: nothing in Shapley's axioms hints at the marginality principle, of long tradition in economic theory.

In the clarification of this puzzle, Young (1985) provides a key answer. He formulates the marginality principle as an axiom, i.e., that the solution should be a function of players' marginal contributions to coalitions. He drops additivity and null player as requirements. The result is that the only solution satisfying efficiency, anonymity and marginality is the Shapley value.

The presence of externalities across coalitions is an important feature in many applications. A few examples spring to mind. In an oligopolistic market, the profit of a cartel depends on the level of cooperation among the competing firms. The power of a political alliance usually depends on the level of coordination among competing parties. The benefit of an agent that refuses to participate in the production of a public good depends on the level of cooperation of the other agents (free-riding effect), and so on.

In this paper we are concerned with the understanding of the marginality principle, when it is combined with externalities across coalitions. The model we shall employ is that of games in partition function form, in which the worth of a coalition S may vary with how the players not in S get organized. In the model, $v(S, \Pi)$ is the worth of S when the coalition structure is Π, S being an element of Π. In defining player i's marginal contribution to coalition S, it is crucial to describe what happens after i leaves S. Suppose i plans to join T, another coalition in Π. The *total effect* on S of i's move is the difference $v(S, \Pi) - v(S \setminus \{i\}, \{S \setminus \{i\}, T \cup \{i\}\} \cup \Pi_{-S,-T})$. This effect can be decomposed into two. First, there is an *intrinsic marginal contribution effect* associated with i leaving S but before joining T, i.e., $v(S, \Pi) - v(S \setminus \{i\}, \{S \setminus \{i\}, \{i\}\} \cup \Pi_{-S})$. And second, there is an *externality effect*, which stems from the change in the worth of $S \setminus \{i\}$ when i, instead of remaining alone, joins T, i.e., the difference $v(S \setminus \{i\}, \{S \setminus \{i\}, \{i\}\} \cup \Pi_{-S}) - v(S \setminus \{i\}, \{S \setminus \{i\}, T \cup \{i\}\} \cup \Pi_{-S,-T})$. (Note how this latter difference is *not* a "partial derivative," a marginal contribution of player i to coalition S.) Our results follow from insisting on this decomposition.

In the first part of the paper, we shall impose that the grand coalition N forms and players' payoffs must add up to $v(N)$. Then we investigate the implications of anonymity, together with a weak version of marginality. According to this, the solution may depend on the total effect of the intrinsic marginal contribution and the externality effects. We find the first noteworthy difference with respect to the case of no-externalities, because in our larger domain these axioms do not even rule out non-linear solutions.

As a result, we strenghthen the weak version of marginality, and we do so in two ways. First, we require monotonicity, i.e., a player's payoff should be increasing in the total effect of his intrinsic marginal contribution and externality effects. Then, we are able to establish upper and lower bounds to each player's payoff in the game. And second, complementing this result, we require a marginality axiom, according to which a player's payoff should depend on the vector of intrinsic marginal contributions, not on the externality effect. The result is a characterization of an "externality-free" value on the basis of efficiency, anonymity and marginality. In a second characterization result, this solution is also obtained using a system of axioms much like the original one due to Shapley.

Thus, it is apparent that using the axioms behind the Shapley value in the larger domain leads to the "externality-free" value. Obviously, in dealing with externalities, this is not enough, and this is why we insist on the combination of results. That is, we view the combination of both kinds of results -the obtention of bounds around the "externality-free" value- as a way to understand how externalities might benefit or punish a player in a context where these normative principles are in place. In effect, the two results together provide a range for acceptable Pigouvian-like transfers (taxes or subsidies among players) when efficiency is accompanied by our other normative desiderata.

The second part of our study does not impose efficiency for the grand coalition. We consider payoff configuration solutions, in which a payoff vector is assigned to each coalition structure. By requiring the principle of balanced contributions (Myerson (1980)) or invoking the notion of potential (Hart and Mas-Colell (1989)), we are able to characterize a unique payoff configuration for each game. This corresponds to a Shapley value of a game in characteristic function form and coalition structures (Aumann and Dreze (1974)).

Based on this characterization, we propose a simple game of coalition formation to answer the question of which coalitions will form. This game yields a unique equilibrium prediction in almost every partition function. Furthermore, its simple structure and sharp prediction should make it useful in applications.

The first papers that proposed value concepts for games with externalities were Myerson (1977) and Bolger (1989). More recently, Fujinaka (2004), Macho-Stadler, Perez-Castrillo and Wettstein (2004) and Pham Do and Norde (2002) also apply the axiomatic approach to the problem, and obtain interesting conclusions. We discuss the relationship between the works of these authors and ours in the next sections. Finally, the coalition formation part of our paper continues a rich literature on these issues (e.g., Bloch (1996, 2002), Ray and Vohra (1999, 2001), Maskin (2003)).

2 Definitions

Let N be the finite set of *players*. A *coalition* is a subset of N. We shall use lower case letters to denote the number of players in a coalition ($s = |S|$, $n = |N|$, etc.). A *partition* is a set $\Pi = \{(S_k)_{k=1}^K\}$ $(1 \leq K \leq n)$ of disjoint coalitions that cover N, i.e. $S_i \cap S_j = \emptyset$, for each $1 \leq i < j \leq K$, and $N = \cup_{k=1}^K S_k$. By convention, $\{\emptyset\} \in \Pi$ for every partition Π. Elements of a partition are called *atoms*. A partition Π' is *finer* than a partition Π if each atom of Π' is included in an atom of Π: if $S' \in \Pi'$, then $S' \subseteq S$ for some $S \in \Pi$. We will say equivalently that Π is *coarser* than Π'. An *embedded coalition* is a pair (S, Π) where Π is a partition and S is an atom of Π. EC denotes the set of embedded coalitions. If S is a coalition and i is a member of S, then S_{-i} (resp. S_{+i}) denotes the set $S \setminus \{i\}$ (resp. $S \cup \{i\}$). Similarly, if Π is a partition and S is an atom of Π, then Π_{-S} denotes the partition $\Pi \setminus \{S\}$ of the set $N \setminus S$.

A *game in partition function form* (Thrall and Lucas (1963)) is a function v that assigns to every embedded coalition (S, Π) a real number $v(S, \Pi)$. We assume that $v(\{\emptyset\}, \Pi) = 0$ for all Π. There are *positive externalities* if $v(S, \Pi) \geq v(S, \Pi')$ for each pair of embedded coalitions (S, Π) and (S, Π') such that Π is coarser than Π'. There are *negative externalities* if $v(S, \Pi) \leq v(S, \Pi')$ for each pair of embedded coalitions (S, Π) and (S, Π') such that Π is coarser than Π'. There are *no externalities* if $v(S, \Pi) = v(S, \Pi')$ for each pair of embedded coalitions (S, Π) and (S, Π'). In the latter case, a partition function is called a characteristic function.

The game v is *superadditive* if $v(S, \Pi) + v(T, \Pi) \leq v(S \cup T, \Pi')$, for every pair (S, Π) and (T, Π) of disjoint embedded coalitions, where $\Pi_{-S,-T} = \Pi'_{-S,-T}$. (We denote $\Pi_{-(S \cup T)}$ by $\Pi_{-S,-T}$.)

A *value* is a function σ that assigns to every game v in partition function form a unique utility vector $\sigma(v) \in \mathbb{R}^N$. Shapley (1953) defined and axiomatized a value on the class of games in characteristic function form:

$$\mathrm{Sh}_i(v) := \sum_{S \subseteq N \text{ s.t. } i \in S} \frac{(s-1)!(n-s)!}{n!} [v(S) - v(S_{-i})]$$

for each player $i \in N$ and each game v in characteristic function form.

3 Weak Marginality

Based on the marginality principle, Young (1985) proposes a beautiful axiomatization of the Shapley value for games in characteristic function. We shall explore the implications of marginality, together with other basic axioms, for the class of games in partition function form. The first two axioms that we shall impose are hardly controversial.

Anonymity *Let π be a permutation of N and let v be a game in partition function form. Then $\sigma(\pi(v)) = \pi(\sigma(v))$, where $\pi(v)(S, \Pi) = v(\pi(S), \{\pi(T) | T \in \Pi\})$ for each embedded coalition (S, Π) and $\pi(x)_i = x_{\pi(i)}$ for each $x \in \mathbb{R}^N$ and each $i \in N$.*

Efficiency *Let v be a game in partition function form. Then $\sum_{i \in N} \sigma_i(v) = v(N)$.*

Anonymity means that players' payoffs do not depend on their names. Efficiency means that the value must be feasible and must exhaust all the benefits from cooperation. Our axiom of efficiency admits two interpretations. Either the game is superadditive and we require the value to be overall efficient, or the grand coalition has to form for some exogenous reason and we require that the outcome yield payoffs that exhaust $v(N)$.

Next, we turn to our discussion of marginality, central in our work. The marginal contribution of a player i within a coalition S is defined, for games in characteristic function form, as the loss incurred by the other members of S if i leaves the group. This number could depend on the organization of the players not in S when there are externalities. It is natural therefore to define the marginal contribution of a player within each embedded coalition.

To begin, one may consider the general case where a player may join another coalition after leaving S. That is, we introduce a (weak) version of the marginality axiom that coincides with Young's concept of marginal contributions on the class of games without externalities.

Weak Version of Marginality *Let v and v' be two games in partition function form. If*

$$v(S,\Pi)-v(S_{-i},\{S_{-i},T_{+i}\}\cup\Pi_{-S,-T}) = v'(S,\Pi)-v'(S_{-i},\{S_{-i},T_{+i}\}\cup\Pi_{-S,-T})$$

for each embedded coalition (S,Π) such that $i\in S$ and each atom T of Π different from S, then $\sigma_i(v)=\sigma_i(v')$.

For instance, if there are three players, then player i's payoff should depend only on the following seven real numbers:

$$\begin{aligned}
A_i(v) &= v(N,\{N\}) - v(\{j,k\},\{\{i\},\{j,k\}\}),\\
B_i(v) &= v(\{i,j\},\{\{i,j\},\{k\}\}) - v(\{j\},\{\{j\},\{i,k\}\}),\\
C_i(v) &= v(\{i,j\},\{\{i,j\},\{k\}\}) - v(\{j\},\{\{i\},\{j\},\{k\}\}),\\
D_i(v) &= v(\{i,k\},\{\{i,k\},\{j\}\}) - v(\{k\},\{\{k\},\{i,j\}\}),\\
E_i(v) &= v(\{i,k\},\{\{i,k\},\{j\}\}) - v(\{k\},\{\{i\},\{j\},\{k\}\}),\\
F_i(v) &= v(\{i\},\{\{i\},\{j,k\}\}),\\
G_i(v) &= v(\{i\},\{\{i\},\{j\},\{k\}\}).
\end{aligned}$$

There is no hope to get a characterization result of a value with this weak notion of marginality. To see this, consider the following examples:

Example 1 *The value σ^α, defined by $\sigma_i^\alpha(v) := \frac{1}{3}A_i(v) + \frac{1}{6}(\alpha B_i(v) + (1-\alpha)C_i(v)) + \frac{1}{6}(\alpha D_i(v) + (1-\alpha)E_i(v)) + \frac{1}{3}(\alpha F_i(v) + (1-\alpha)G_i(v))$, satisfies the anonymity and the efficiency axioms, as well as the weak version of marginality, for every $\alpha\in\mathbb{R}$. The values σ^α are instances of the average approach characterized by Macho-Stadler et al. (2004), as they coincide with the Shapley value of a fictitious game v^α in characteristic function form, where $v^\alpha(\{i\}) = \alpha v(\{i\},\{\{i\},\{j,k\}\}) + (1-\alpha)v(\{i\},\{\{i\},\{j\},\{k\}\})$.*

In addition, and what is perhaps more surprising, a large class of non-linear values satisfy the three axioms. (Recall that they imply linearity in the domain of games in characteristic function form.) In this sense, our approach differs substantially from Fujinaka's (2004). He proposes several versions of marginality, whereby a marginal contribution is constructed as a weighted linear average of the marginal contributions over different coalition structures. This assumption already builds linearity in Fujinaka's result.

Example 2 *Let $m : \mathbb{R} \to \mathbb{R}$ be any function. Then the value $\sigma^{\alpha,m}$, defined by $\sigma_i^{\alpha,m}(v) := \sigma_i^{\alpha}(v) + m(F_i(v) - G_i(v)) - \frac{m(C_i(v)-B_i(v))+m(E_i(v)-D_i(v))}{2}$, also satisfies the three axioms. Observe that the differences $C_i(v) - B_i(v)$, $E_i(v) - D_i(v)$, and $F_i(v) - G_i(v)$ measure the externality that the agents face. The function m transforms the externality that a player faces into a transfer paid equally by the two other players. The value $\sigma^{\alpha,m}$ is then obtained by adding to σ^{α} the net transfer that each player receives.*

Given these examples, we propose to follow two alternative paths. First, we shall strengthen the weak marginality axiom into a monotonicity property. Second, we shall look more closely at the notion of "marginal contributions" to propose an alternative marginality axiom. We undertake each of the alternatives in the next two sections.

4 Monotonicity

This section investigates what happens when, in addition to requiring efficiency and anonymity, weak marginality is strenghthened to the following monotonicity axiom. The result will be the derivation of useful bounds to the payoff of each player.

Monotonicity *Let v and v' be two games in partition function form. If*

$$v(S,\Pi) - v(S_{-i}, \{S_{-i}, T_{+i}\} \cup \Pi_{-S,-T}) \geq v'(S,\Pi) - v'(S_{-i}, \{S_{-i}, T_{+i}\} \cup \Pi_{-S,-T})$$

for each embedded coalition (S,Π) such that $i \in S$ and each atom T of Π different from S, then $\sigma_i(v) \geq \sigma_i(v')$.

In words, if in a game the vector of marginal contributions of a player to the different coalitions, for any organization of the complement, dominates coordinate by coordinate that of a second game, the value must pay this player more in the first game. For instance, the value σ^{α} is monotonic, for each $\alpha \in [0,1]$.

First, we point out in the following example that monotonicity combined with anonymity and efficiency does not imply additivity either:

Example 3 *The value $\sigma^{\alpha,m}$ from Example 2 is monotonic if $\alpha = 1/2$, $m(x) = x^2$ if $|x| \leq 1/12$, and $m(x) = (1/12)^2$ if $|x| \geq 1/12$.*

We may nevertheless bound each player's payoff from below and from above. For each player i, let $\underline{v}^i$ and $\bar{v}^i$ be two fictitious games in characteristic function form defined as follows:

$$\underline{v}^i(S) := \begin{cases} \min_\Pi\{v(S,\Pi)|S \in \Pi\} & \text{if } i \in S \\ \max_\Pi\{v(S,\Pi)|S \in \Pi\} & \text{if } i \in N \setminus S \end{cases}$$

$$\bar{v}^i(S) := \begin{cases} \max_\Pi\{v(S,\Pi)|S \in \Pi\} & \text{if } i \in S \\ \min_\Pi\{v(S,\Pi)|S \in \Pi\} & \text{if } i \in N \setminus S \end{cases}$$

for each coalition S.

Thus, the game in characteristic function form $\underline{v}^i$ represents pessimistic expectations from the point of view of player i, who assumes that, whenever a coalition that contains him cooperates, the complement organizes itself to minimize the coalition's worth. Exactly the opposite happens for coalitions that exclude him. In contrast, the game $\bar{v}^i$ is built on player i's most optimistic expectations.

For example, if there are negative externalities, then

$$\underline{v}^i(S) := \begin{cases} v(S,\{S, N \setminus S\}) & \text{if } i \in S \\ v(S,\{S,\{j\}_{j \in N\setminus S}\}) & \text{if } i \in N \setminus S, \end{cases}$$

and

$$\bar{v}^i(S) := \begin{cases} v(S,\{S,\{j\}_{j \in N\setminus S}\}) & \text{if } i \in S \\ v(S,\{S, N \setminus S\}) & \text{if } i \in N \setminus S. \end{cases}$$

And if there are positive externalities, then

$$\underline{v}^i(S) := \begin{cases} v(S,\{S,\{j\}_{j \in N\setminus S}\}) & \text{if } i \in S \\ v(S,\{S, N \setminus S\}) & \text{if } i \in N \setminus S, \end{cases}$$

and

$$\bar{v}^i(S) := \begin{cases} v(S,\{S, N \setminus S\}) & \text{if } i \in S \\ v(S,\{S,\{j\}_{j \in N\setminus S}\}) & \text{if } i \in N \setminus S. \end{cases}$$

Of course, $\underline{v}^i = \bar{v}^i$ if there are no externalities.

We are now ready to state our first result:

Proposition 1 *If σ is a value that satisfies anonymity, efficiency and monotonicity, then for each $i \in N$,*

$$\sigma_i(v) \in [Sh_i(\underline{v}^i), Sh_i(\bar{v}^i)].$$

Proof: Let V be the vector space of games in partition function form with no externalities. We define an isomorphism between V and the set of games in characteristic function form: $\gamma(v)(S) := v(S,\Pi)$ for each game $v \in V$ and each coalition S; $\gamma^{-1}(v')(S,\Pi) := v'(S)$ for each game v' in characteristic function and each embedded coalition (S,Π).

The value $\hat{\sigma} := \sigma \circ \gamma^{-1}$ satisfies Young's (1985) axioms on the space of games in characteristic function form. Hence, $\sigma \circ \gamma^{-1} = \mathrm{Sh}$, or $\sigma = \mathrm{Sh} \circ \gamma$ for every $v \in V$.

Now let v be an arbitrary game in partition function form, and let $\nu^i \in V$ be the game defined as follows:

$$\nu^i(S,\Pi) := \begin{cases} \max_{\Pi'}\{v(S,\Pi')|S \in \Pi'\} & \text{if } i \in S \\ \min_{\Pi'}\{v(S,\Pi')|S \in \Pi'\} & \text{if } i \in N \setminus S \end{cases}$$

for each embedded coalition (S,Π). Monotonicity implies that $\sigma_i(v) \leq \sigma_i(\nu^i) = \mathrm{Sh}_i(\gamma(\nu^i)) = \mathrm{Sh}_i(\bar{v}^i)$. A similar reasoning applies to establish the other inequality. $\square$

Thus, anonymity, efficiency and monotonicity imply that each player i's payoff must be bounded between his Shapley value payoffs in his most pessimistic and most optimistic games in characteristic function.

5 Marginality

An alternative route to monotonicity is to strenghthen the weak version of marginality into another marginality axiom. To do this, it will be instructive to look closer at the concept of marginal contribution in contexts with externalities.

Consider player i and an embedded coalition (S,Π) with $i \in S$. Suppose player i leaves coalition S and joins coalition $T \in \Pi$, $T \neq S$. One can view this as a two-step process. In the first instance, player i simply leaves S and, at least for a while, he is alone, which means that for the moment the coalition structure is $\{S_{-i}, \{i\}\} \cup \Pi_{-S}$. At this point, coalition S_{-i} feels the loss of player i's *marginal contribution*, i.e.,

$$v(S,\Pi) - v(S_{-i}, \{S_{-i}, \{i\}\} \cup \Pi_{-S}).$$

In the second step, player i joins coalition $T \in \Pi_{-S}$, and then S_{-i} is further affected, but not because of a marginal contribution from player i. Rather, it is affected because of the corresponding *externalities* created by this merger, i.e.,

$$v(S_{-i}, \{S_{-i}, \{i\}\} \cup \Pi_{-S}) - v(S_{-i}, \{S_{-i}, T_{+i}\} \cup \Pi_{-S,-T}).$$

If one views this as an important distinction, one should reserve the term (intrinsic) *marginal contribution* to the former difference. We shall do this in the sequel.

Formally, let i be a player and let (S,Π) be an embedded coalition such that $i \in S$. Then the (intrinsic) *marginal contribution of* i *to* (S,Π) is given by

$$mc_{(i,S,\Pi)}(v) = v(S,\Pi) - v(S_{-i}, \{S_{-i}, \{i\}\} \cup \Pi_{-S})$$

for each game v in partition function. Player i's *vector of intrinsic marginal contributions* is obtained by varying (S,Π): $mc_i(v) = (mc_{(i,S,\Pi)})_{(S,\Pi) \in EC \wedge i \in S}$. Here is the formal statement of the new marginality axiom for games in partition function form (note that it also reduces to Young's if applied to games in characteristic form).

Marginality *Let i be a player and let v and v' be two games in partition function form. If $mc_i(v) = mc_i(v')$, then $\sigma_i(v) = \sigma_i(v')$.*

Consider now the following extension σ^* of the Shapley value to the class of games in partition function form:

$$\sigma_i^*(v) := \mathrm{Sh}_i(\hat{v})$$

for each player $i \in N$ and each game v in partition function form, where $\hat{v}$ is the fictitious game in characteristic function form defined as follows:

$$\hat{v}(S) := v(S, \{S, \{j\}_{j \in N \setminus S}\})$$

for each coalition S.

We can call this the "externality-free" value, and we shall discuss it below. Our second main result follows.

Proposition 2 *σ^* is the unique value satisfying anonymity, efficiency and marginality.*

Proof: The set of games in partition function form is a vector space. We define a basis of this space to prove the proposition. Let (S, Π) be an embedded coalition, where S is non-empty. Then $e_{(S,\Pi)}$ is the game in partition function form defined as follows:

$$e_{(S,\Pi)}(S', \Pi') = \begin{cases} 1 & \text{if } S \subseteq S' \text{ and } (\forall T' \in \Pi' \setminus \{S'\})(\exists T \in \Pi) : T' \subseteq T, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 1 *The collection of vectors $(e_{(S,\Pi)})_{(S,\Pi) \in EC}$ constitutes a basis of the space of games in partition function form.*

Proof of Lemma 1: The number of vectors in the collection equals the dimension of the space. It is therefore enough to show that they are linearly independent. Let $(\alpha_{(S,\Pi)})_{(S,\Pi) \in EC}$ be a collection of real numbers such that

$$\sum_{(S,\Pi) \in EC} \alpha(S, \Pi) e_{(S,\Pi)} = 0. \tag{1}$$

We have to show that $\alpha(S, \Pi) = 0$ for each $(S, \Pi) \in EC$. Suppose on the contrary that there exists $(S, \Pi) \in EC$ such that $\alpha(S, \Pi) \neq 0$. Let (S^*, Π^*) be an embedded coalition such that:

1. $\alpha(S^*, \Pi^*) \neq 0$;
2. there does not exist an embedded coalition (S, Π) with $\alpha(S, \Pi) \neq 0$ and $S \subsetneq S^*$; and
3. there does not exist an embedded coalition (S^*, Π) with $\alpha(S^*, \Pi) \neq 0$ and Π different and coarser than Π^*.

Observe that $\alpha(S,\Pi)=0$ if S is a strict subset of S^*. Also, $e_{(S,\Pi)}(S^*,\Pi^*)=0$ if S is not included in S^*. Hence,

$$[\sum_{(S,\Pi)\in EC}\alpha(S,\Pi)e_{(S,\Pi)}](S^*,\Pi^*)=[\sum_{(S^*,\Pi)\in EC}\alpha(S^*,\Pi)e_{(S^*,\Pi)}](S^*,\Pi^*).$$

Now observe that $\alpha(S^*,\Pi)=0$ if Π is different and coarser than Π^*. Also, $e_{(S^*,\Pi)}(S^*,\Pi^*)=0$ if Π is not coarser than Π^*. Hence,

$$[\sum_{(S^*,\Pi)\in EC}\alpha(S^*,\Pi)e_{(S^*,\Pi)}](S^*,\Pi^*)=\alpha(S^*,\Pi^*).$$

Equation (1) then implies that $\alpha(S^*,\Pi^*)=0$, a contradiction. $\Box$

Now we continue with the proof of Proposition 2. Given the properties of the Shapley value, it is easy to check that σ^* satisfies the three axioms.

We prove uniqueness. Let σ be a value satisfying the three axioms. We show that $\sigma=\sigma^*$ by induction on the number of non-zero terms appearing in the decomposition of the game in the basis.

Suppose first that there is just one term, i.e. $v=\alpha e_{(S,\Pi)}$ for some $\alpha\in\mathbb{R}$ and some $(S,\Pi)\in EC$. We now check that

$$\sigma_i^*(e_{(S,\Pi)})=\begin{cases}\alpha/s & \text{if } i\in S\\ 0 & \text{if } i\in N\setminus S.\end{cases}$$

To do this, let $i\in N\setminus S$ and let (S',Π') be any embedded coalition. The two following statements are equivalent:

1. $S\subseteq S'$ and $(\forall T'\in\Pi'_{-S'})(\exists T\in\Pi):T'\subseteq T$;

2. $S\subseteq S'\setminus\{i\}$ and $(\forall T'\in\{\{i\},\Pi'_{-S'}\})(\exists T\in\Pi):T'\subseteq T$.

Hence $e_{(S,\Pi)}(S',\Pi')=1$ if and only if $e_{(S,\Pi)}(S'_{-i},\{S'_{-i},\{i\},\Pi'_{-S'}\})=1$. So, $mc_i(v)=mc_i(v^0)$, where v^0 is the null game (i.e., $v^0(S,\Pi)=0$ for each $(S,\Pi)\in EC$). Marginality implies that $\sigma_i(v)=\sigma_i(v^0)$. On the other hand, anonymity and efficiency imply that $\sigma_i(v^0)=0$. Hence $\sigma_i(v)=0$.

Let now i and j be two members of S. Let π be the permutation that exchanges i with j. Then $\pi(e_{(S,\Pi)})=e_{(S,\Pi)}$ and $\sigma_i(v)=\sigma_j(v)$, as σ is anonymous. Finally, the efficiency axiom implies that $\sigma_i(v)=\alpha/s$ for each member i of S.

Suppose now that we have proved the result for all the games that have at most k non-zero terms when decomposed in the basis and let

$$v=\sum_{(S,\Pi)\in EC}\alpha(S,\Pi)e_{(S,\Pi)}$$

be a game with exactly $k+1$ non-zero coefficients. Let S^* be the intersection of the coalitions S for which there exists a partition Π such that $\alpha(S,\Pi)$ is different from zero.

If $i \in N \setminus S^*$, then player i's marginal contribution vector in the game v coincides with his marginal contribution vector in the game

$$v' = \sum_{(S,\Pi)\in EC \text{ s.t. } i\in S} \alpha(S,\Pi) e_{(S,\Pi)}.$$

Marginality implies that $\sigma_i(v) = \sigma_i(v')$. Note that in the game v' the number of terms for which $\alpha(S,\Pi) \neq 0$ is at most k. Then, by the induction hypothesis, $\sigma_i(v') = \sigma_i^*(v')$. Since the value σ^* satisfies the marginality axiom as well, we conclude that $\sigma_i(v) = \sigma_i^*(v)$.

On the other hand, anonymity implies that for each player $i, j \in S^*$, $\sigma_i(v) = \sigma_j(v)$ and $\sigma_i^*(v) = \sigma_j^*(v)$. Since both σ and σ^* satisfy efficiency, for all $i \in S^*$, $\sigma_i(v) = \sigma_i^*(v)$. The proof of Proposition 2 is now complete. $\square$

Proposition 2 is not a trivial variation on Young's (1985) original theorem. Some information present in the partition function has to be discarded, as a consequence of the combination of the three axioms. To gain an intuition for this, consider first a three-player game. The result tells us that player 1's payoff does not depend on $x = v(\{1\}, \{\{1\}, \{2,3\}\})$. This does not follow from any of the three axioms taken separately (in particular, by marginality it could depend on x, if $S = \{1\}$ and $\Pi_{-S} = \{\{2,3\}\}$ in the definition of the axiom). Instead, the reasoning goes as follows. Players 2 and 3's payoffs do not depend on x according to the marginality axiom. The efficiency axiom then implies that player 1's payoff cannot depend on x either.

To gain more intuition, let us continue with this heuristic argument and see how it would apply beyond 3-player games. Consider a four-player game. In principle, player 1's payoff could depend on fifteen numbers according to the marginality axiom. Only eight of these numbers are actually relevant to compute σ^*. Let us show for instance why player 1's payoff cannot depend on $y = v(\{1\}, \{\{1\}, \{2,3\}, \{4\}\})$. Marginality implies that, apart from player 1's, only the payoff of player 4 could depend on y, or more precisely on $z - y$, where $z = v(\{1,4\}, \{\{1,4\}, \{2,3\}\})$. Marginality implies also that the payoffs of players 2 and 3 do not depend on z. On the other hand, as we know from the proposition, the three axioms together imply that the solution must be an anonymous and additive function. Thus, the payoffs of players 1 and 4 depend identically on z (if z increases, both payoffs to players 1 and 4 also increase). Hence, efficiency implies that player 4's payoff cannot depend on z, and therefore not on y either. The payoffs of players 2 and 3 do not depend on y either by the marginality axiom. Hence efficiency implies that player 1's payoff cannot depend on y.

We regard the "externality-free" value σ^* as a fair compromise that takes into account the pure or intrinsic *marginal contributions* of players to coalitions, stripped down from externality components. The range of payoffs identified in Proposition 1 for each player captures how the externalities affect his payoff, when one still requires efficiency, anonymity and monotonicity. Thus, the size of the difference $\mathrm{Sh}_i(\bar{v}^i) - \sigma_i^*$ expresses the maximum "subsidy" or benefit to player

i, associated with externalities that favor him, and the difference $\sigma_i^* - \mathrm{Sh}_i(\underline{v}^i)$ represents how much i can be "taxed" or suffer, due to harmful externalities, in a value that is still faithful to efficiency, anonymity and a principle that heavily relies on marginal considerations.

The first justification of the efficiency axiom was referring to superadditive games. Proposition 2 remains true even if we restrict our attention to this subclass of games: σ^* is the unique value satisfying the three axioms (anonymity, efficiency and marginality) on the class of superadditive games. We need to adapt the proof, as in Young (1985, page 71), since the game v' appearing in the proof is not necessarily superadditive. Let σ be a value that satisfies the three axioms on the class of superadditive games and let v be a superadditive game. Let

$$\begin{cases} \bar{\alpha}(s) := \max_{(S,\Pi)\in EC \text{ s.t. } \#S=s} \alpha(S,\Pi), \text{ for each } 1 \leq s \leq n, \\ u := \sum_{(S,\Pi)\in EC} \bar{\alpha}(s) e_{(S,\Pi)}, \\ \alpha'(S,\Pi) := \bar{\alpha}(s) - \alpha(S,\Pi), \\ w := \sum_{(S,\Pi)\in EC} \alpha'(S,\Pi) e_{(S,\Pi)}. \end{cases}$$

Notice that $\alpha'(S,\Pi) \geq 0$ for each $(S,\Pi) \in EC$. Also, the sum of superadditive games is a superadditive game, and each game in the basis is superadditive. Hence w and $u = v + w$ are superadditive. On the other hand, u is symmetric, in the sense that $u(S,\Pi) = u(S',\Pi')$ for each pair of embedded coalitions (S,Π) and (S',Π') such that S and S' have the same cardinality. Anonymity and efficiency imply that $\sigma(u) = \sigma^*(u)$. It is then easy to apply the induction argument of the previous proof to the number of non-zero terms appearing in the decomposition of w. The key difference in using this alternative decomposition of v is that the game obtained by deleting any number of terms $\alpha'(S,\Pi)e_{(S,\Pi)}$ is still superadditive.

One can take a bargaining approach to understand the Shapley value. This was done for example in Hart and Mas-Colell (1996) or Perez-Castrillo and Wettstein (2001) for games in characteristic function form. It is not difficult to see that we obtain σ^* if we apply these procedures to superadditive games in partition function form. Indeed, it is assumed in these two papers that, when a proposal is rejected, the rejected proposer goes off by himself and does not form a coalition with anyone else. We leave the details to the interested reader. Other rules concerning the rejected proposers would lead to values that treat externalities differently (see Macho-Stadler et al. (2005)).

6 Additivity

A natural adaptation of Shapley's original axiomatic system also leads to σ^*. Let i be a player and let v be a game in partition function form. We say that player i is *null* if his marginal contribution to any embedded coalition is nihil: $mc_i(v) = 0$. Here are two additional key axioms that can be used to obtain an alternative characterization of σ^*.

Null Player *Let $i \in N$ and let v be a game in partition function form. If player i is null, then $\sigma_i(v) = 0$.*

Additivity *Let v and w be two games in partition function form. Then, $\sigma(v + w) = \sigma(v) + \sigma(w)$.*

A null player must receive a zero payoff, according to the null player axiom. Additivity essentially amounts to the linearity of the value. It expresses a form of mathematical simplicity by requiring a strong specific functional form.

Proposition 3 *σ^* is the unique value satisfying anonymity, efficiency, null player and additivity.*

Proof: Given the properties of the Shapley value, it is easy to check that σ^* satisfies the four axioms.

Let σ be a value satisfying the four axioms. Let v^0 be the null game: $v^0(S, \Pi) = 0$ for each embedded coalition (S, Π). By the additivity and the null player axioms, $\sigma(v) + \sigma(-v) = \sigma(v^0) = 0$ for each game v in partition function. Combining this observation with the additivity axiom, we conclude that σ is linear. It is therefore sufficient to prove that σ coincides with σ^* on the basis e defined in the previous section. This was already done in the first step of the induction proof for proposition 2. $\square$

Proposition 3 is equivalent to theorem 2 of Pham Do and Norde (2002). It is not difficult to check that their solution based on the random arrival principle coincides with σ^*. There seems to be a problem in their proof, though. Indeed, they intend to apply similar arguments with an alternative basis. Let (S, Π) be an embedded coalition. Then $e'_{(S,\Pi)}$ is the game in partition function form defined as follows:

$$e'_{(S,\Pi)}(S', \Pi') = \begin{cases} 1 & \text{if } S \subseteq S' \text{ and } (\forall T' \in \Pi' \setminus \{S'\})(\exists T \in \Pi) : T' = T \cap (N \setminus S'), \\ 0 & \text{otherwise.} \end{cases}$$

Pham Do and Norde claim that players outside S are null in $e'_{(S,\Pi)}$, but this is incorrect. Consider for instance the game $v' = e'_{(\{1\},\{\{1\},\{2,3\}\})}$ (cf. τ_9 in their example 4.2). Then $v'(\{1, 2\}, \{\{1, 2\}, \{3\}\}) = 1$, while $v'(\{1\}, \{\{1\}, \{2\}, \{3\}\}) = 0$. Hence, one cannot say that player 2 is null in v'.

Proposition 3 also remains true even if we restrict ourselves to the subclass of superadditive games: σ^* is the unique value satisfying the four axioms (anonymity, efficiency, null player and additivity) on the class of superadditive games. Let σ be a value that satisfies the four axioms on the class of superadditive games. Let v be a superadditive game in partition function form. Let $(\alpha(S, \Pi))_{(S,\Pi)\in EC}$ be the vector of real numbers such that $v = \sum_{(S,\Pi)\in EC} \alpha(S, \Pi) e_{(S,\Pi)}$. Then

$$v + \sum_{\{(S,\Pi)|\alpha(S,\Pi)<0\}} |\alpha(S, \Pi)| e_{(S,\Pi)} = \sum_{\{(S,\Pi)|\alpha(S,\Pi)\geq 0\}} \alpha(S, \Pi) e_{(S,\Pi)}.$$

Notice that each game in the basis e is superadditive. Additivity implies that

$$\sigma(v) + \sum_{\{(S,\Pi)|\alpha(S,\Pi)<0\}} |\alpha(S,\Pi)|\sigma(e_{(S,\Pi)}) = \sum_{\{(S,\Pi)|\alpha(S,\Pi)\geq 0\}} \alpha(S,\Pi)\sigma(e_{(S,\Pi)}).$$

Null player, anonymity, and efficiency still imply that σ coincides with σ^* on the basis. Hence,

$$\sigma(v) + \sum_{\{(S,\Pi)|\alpha(S,\Pi)<0\}} |\alpha(S,\Pi)|\sigma^*(e_{(S,\Pi)}) = \sum_{\{(S,\Pi)|\alpha(S,\Pi)\geq 0\}} \alpha(S,\Pi)\sigma^*(e_{(S,\Pi)}).$$

The linearity of σ^* implies that $\sigma(v) = \sigma^*(v)$.

A player is null if his vector of marginal contributions is nihil. The other notion of marginal contributions, which contained the externality effects and was used to define the weak version of the marginality axiom, leads in turn to a weaker version of the null player axiom. Player i is *null in the strong sense* if

$$v(S,\Pi) - v(S_{-i}, \{S_{-i}, T_{+i}\} \cup \Pi_{-S,-T}) = 0$$

for each embedded coalition (S,Π) such that $i \in S$ and each atom T of Π different from S. Clearly, if player i is null in the strong sense, then he is null. One can use this definition to propose a different null player axiom:

Weak Version of the Null Player Axiom *Let $i \in N$ and let v be a game in partition function form. If player i is null in the strong sense, then $\sigma_i(v) = 0$.*

This is equivalent to the dummy player axiom of Bolger (1989) and Macho-Stadler et al. (2004). They say that player i is dummy if $v(S,\Pi) = v(S',\Pi')$ for each (S,Π) and each (S',Π') that can be deduced from (S,Π) by changing player i's affiliation. This clearly implies that player i is null in our strong sense. The converse is straightforward after proving that $v(S,\Pi) = v(S,\Pi')$ for each pair of embedded coalitions (S,Π) and (S,Π') such that $i \notin S$ and (S,Π') can be deduced from (S,Π) by changing only player i's affiliation. Indeed, if i is null in our strong sense, then $v(S,\Pi) = v(S_{+i}, \{S_{+i}\} \cup \{T_{-i}|T \in \Pi_{-S}\}) = v(S_{+i}, \{S_{+i}\} \cup \{T_{-i}|T \in \Pi'_{-S}\}) = v(S,\Pi')$.

Macho-Stadler et al. (2004, theorem 1) show that any solution that satisfies this version of the null player axiom, as well as the axioms of efficiency, additivity and (a strong version of) anonymity is a Shapley value of a game in characteristic function form that is obtained by performing averages of the partition function. Our "externality-free" value σ^* belongs to this class of solutions. Macho-Stadler et al. also characterize a unique solution by adding an axiom of similar influence that σ^* does not satisfy.

In light of our first discussion concerning even non-linear solutions (recall Example 2), we prefer the axioms of Proposition 2 to those in Proposition 3, even though both lead to the same value. We find marginality more compelling than additivity. It is easier to interpret and justify a restriction on the set of variables required to compute the payoff of the players, than to impose a specific functional form. It is interesting to note in that respect that the non-additive solution defined in Example 3 satisfies the strong symmetry and the similar influence axioms of Macho-Stadler et al. (2004), in addition to satisfying anonymity, efficiency, and the weak version of the marginality axiom. Once again, additivity cannot be justified by the marginality principle that underlies their dummy player axiom (the weak version of marginality), even if one imposes their other requirements.

7 More Examples

At this point it is probably desirable to illustrate how our results can be used in examples. This section does precisely this, by calculating the range of possible payoffs recommended to each player by each of our different sets of axioms. We present two examples of appealing economic content.

Example 4 *This example features prominently in Ray and Vohra (2001) and Maskin (2003). The example describes a typical "free rider" problem created by a public good that can be produced by each two-player coalition. The set of agents is* $N = \{1, 2, 3\}$*, and this is the partition function:*

$$v(N) = 24;$$

$$v(\{1,2\}) = 12;$$
$$v(\{1,3\}) = 13;$$
$$v(\{2,3\}) = 14;$$
$$v(\{i\}, \{\{i\}, \{j,k\}\}) = 9 \quad \textit{for all } i, j, k;$$
$$v(\{i\}, \{\{i\}, \{j\}, \{k\}\}) = 0 \quad \textit{for all } i, j, k.$$

First, we calculate the payoffs corresponding to the value σ^**. This is done by computing the Shapley value of the game in characteristic function form where each individual i's worth is* $v(\{i\}, \{\{i\}, \{j\}, \{k\}\}) = 0$*. Thus,* $\sigma^*(v) = (7.5, 8, 8.5)$*. This is the payoff recommended if the* marginal contributions*, and not the externalities, are taken into account by the solution. Effectively, This corresponds to the Shapley value where the "free rider" effects are ignored.*

Next, let us calculate the range of payoffs compatible with monotonicity. For each player i, the game $\bar{v}^i$ *in characteristic form is such that* $\bar{v}^i(\{i\}) = 9$ *and* $\bar{v}^i(\{j\}) = 0$ *for* $j \neq i$ *(this corresponds to player i's optimistic expectations of being able to "free ride" on the coalition* $\{j, k\}$ *when he is alone, while he "punishes" the other singleton coalitions by never cooperating in any two-player coalition). The reader can check that* $Sh_1(\bar{v}^1) = 10.5$*,* $Sh_2(\bar{v}^2) = 11$ *and* $Sh_3(v^3) = 11.5$*. Each of these numbers is the upper bound to each player's value payoff, when values satisfy monotonicity.*

In contrast, the game $\underline{v}^i$ *yields* $\underline{v}^i(\{i\}) = 0$ *and* $\underline{v}^i(\{j\}) = 9$ *for* $j \neq i$*. Then, one obtains that* $Sh_1(\underline{v}^1) = 4.5$*,* $Sh_2(\underline{v}^2) = 5$ *and* $Sh_3(\underline{v}^3) = 5.5$*. These are the lower bounds to each player's payoff in values that satisfy monotonicity.*

In particular, we learn that no player can appropriate more than three units of surplus over his payoff in σ^*. If one of the players, making use of a strong bargaining power, appropriates all of those three units, an extreme benefit from the positive externality, no other player can benefit with respect to σ^* at any of these values. Similarly, no player can be punished with more than three units of loss with respect to σ^*, and if a player is pushed all the way down to his lower bound, no other player can be paid less than what he gets at σ^*. Considerations like these are useful in order to evaluate how the benefits/costs associated with externalities must be shared, if one insists on our normative principles.

Example 5 *Consider a variant of the previous example, in which player 1 is the only agent capable of free-riding from a two-player coalition, receiving a worth of 9, as before, when coalition $\{2,3\}$ gets together. However,*

$$v(\{2\},\{\{2\},\{1,3\}\}) = v(\{3\},\{\{3\},\{1,2\}\}) = 0.$$

One can easily check that $\sigma^*(v) = (7.5, 8, 8.5)$, as before. However, the reader will see that:

$$\begin{aligned} Sh_1(\bar{v}^1) &= 10.5, \qquad Sh_1(\underline{v}^1) = 7.5; \\ Sh_2(\bar{v}^2) &= 8, \qquad Sh_2(\underline{v}^2) = 6.5; \\ Sh_3(\bar{v}^3) &= 8.5, \qquad Sh_3(\underline{v}^3) = 7. \end{aligned}$$

That is, no monotonic solution σ ever punishes player 1 or rewards 2 and 3 with respect to the payoffs in σ^*.

8 Moving Away from the Grand Coalition: A Balanced Contributions Approach

The objective of the present section is to adapt Myerson's (1980) principle of balanced contributions in order to apply it to games in partition function form.[2] There are two main differences with respect to the paradigm followed in the previous sections. First, the analysis is conducted for a given game instead of deducing the value by comparison of different games. Second, a payoff vector is specified for each coalition structure. So the payoff of each player is determined as a function of the coalitions that form, instead of assuming that the grand coalition always forms.

Formally, a *payoff configuration* pc is a function that associates to every coalition structure Π a payoff vector pc(Π) in $\mathbb{R}^N$. We now turn to axioms on payoff configurations.

Partition efficiency *For a game v in partition function form, a payoff configuration is partition efficient if*

$$\sum_{i \in S} pc_i(\Pi) = v(S, \Pi),$$

[2] Equivalently, the result in this section can also be viewed as a use of the concept of potential (Hart and Mas-Colell (1989)).

for each embedded coalition (S, Π)*.*

Partition efficiency will now replace the axiom of efficiency used in previous sections.

Next, we introduce the principle of balanced contributions, which requires that all the members of an embedded coalition gain equally from the cooperation of the other members of the embedded coalition. More precisely:

Balanced contributions *The payoff configuration pc satisfies the principle of balanced contributions if*

$$pc_i(\Pi) - pc_i(\{S_{-j}, \{j\}\} \cup \Pi_{-S}) = pc_j(\Pi) - pc_j(\{S_{-i}, \{i\}\} \cup \Pi_{-S}),$$

for each embedded coalition (S, Π) *and each pair* (i, j) *of players that belong to* S*.*

In other words, the payoff loss that player i suffers if player j leaves S given the coalition structure Π is equal to the payoff loss that player j suffers if player i leaves S given the coalition structure Π. Notice how the outside options of the players are determined endogenously, thanks to the concept of payoff configuration. We assume, as when we defined intrinsic marginal contributions before, that a player cannot join another coalition after leaving. Despite this assumption, the externalities do play now an important role in the solution, again because of the payoff configuration notion.

Let S be a coalition. A game in characteristic function form defined on S is a function v that associates a real number $v(T)$ to every non-empty subset T of S. The Shapley value is defined over S as follows:

$$\mathrm{Sh}_i(v) := \sum_{T \subseteq S \text{ s.t. } i \in T} \frac{(t-1)!(s-t)!}{s!}[v(T) - v(T_{-i})]$$

for each player $i \in S$.

This is the main result of this section:

Proposition 4 *Let* v *be a game in partition function form. There exists a unique payoff configuration* pc^* *that is partition efficient for* v*, and that satisfies the principle of balanced contributions. Let* i *be a player, let* Π *be a partition and let* S *be the atom that includes* i*. Then* $pc^*_i(\Pi) = Sh_i(\hat{v}_{(S,\Pi)})$*, where* $\hat{v}_{(S,\Pi)}$ *is a fictitious game in characteristic function form defined over* S *as follows:* $\hat{v}_{(S,\Pi)}(T) = v(T, \{T, \{j\}_{j \in S \setminus T}\} \cup \Pi_{-S})$*, for each non-empty subset* T *of* S*. In particular,* $pc^*(\{N\}) = \sigma^*(v)$*.*

Proof: We prove the result by induction on the cardinality s of the atom S of the partition Π to which i belongs. The result follows from the partition efficiency condition if $s = 1$. Suppose that we proved the result for $1 \leq s \leq n-1$ and let us prove it for all the members of an atom S of a partition Π such that $\#S = s + 1$. Let $\Pi'(i)$ be the partition $\{S_{-i}, \{i\}\} \cup \Pi_{-S}$, for each $i \in S$. Notice that there exists a unique vector $x \in \mathbb{R}^S$ such that $\sum_{i \in S} x_i = v(S, \Pi)$ and that satisfies the following system of equations:

$$x_i - \mathrm{pc}_i^*(\Pi'(j)) = x_j - \mathrm{pc}_j^*(\Pi'(i)),$$

for each pair (i, j) of players that belong to S . Indeed, if x and x' are two solutions, then $x_i - x_i' = x_j - x_j'$ for each pair (i, j) of members of S. Thus, $x = x'$, since $\sum_{i \in S} x_i = \sum_{i \in S} x_i' = v(S, \Pi)$.

In addition, the vector $(\mathrm{pc}_i^*(\Pi))_{i \in S}$ is a solution of the system. Indeed, let i and j be two different members of S. Then

$$\mathrm{pc}_i^*(\Pi) - \mathrm{pc}_j^*(\Pi) = \mathrm{Sh}_i(\hat{v}_{(S,\Pi)}) - \mathrm{Sh}_j(\hat{v}_{(S,\Pi)}).$$

Using the Shapley formula and rearranging the terms, we get:

$$\begin{aligned}\mathrm{pc}_i^*(\Pi) - \mathrm{pc}_j^*(\Pi) = {} & \textstyle\sum_{T \subseteq S_{-j} \text{ s.t. } i \in T} (\gamma(s,t) + \delta(s,t)) \hat{v}_{(S,\Pi)}(T) \\ & - \textstyle\sum_{T \subseteq S_{-i} \text{ s.t. } j \in T} (\gamma(s,t) + \delta(s,t)) \hat{v}_{(S,\Pi)}(T),\end{aligned}$$

where

$$\gamma(s,t) = \frac{(t-1)!(s-t)!}{s!}$$

and

$$\delta(s,t) = \frac{t!(s-t-1)!}{s!}.$$

On the other hand, $\mathrm{pc}_i^*(\Pi'(j)) - \mathrm{pc}_j^*(\Pi'(i))$ may be rewritten as

$$\sum_{T \subseteq S_{-j} \text{ s.t. } i \in T} \epsilon(s,t) \hat{v}_{(S_{-j}, \Pi'(j))}(T) - \sum_{T \subseteq S_{-i} \text{ s.t. } j \in T} \epsilon(s,t) \hat{v}_{(S_{-i}, \Pi'(i))}(T),$$

where

$$\epsilon(s,t) = \frac{(t-1)!(s-t-1)!}{(s-1)!}.$$

The result then follows from the fact that $\gamma(s,t) + \delta(s,t)) = \epsilon(s,t)$, $\hat{v}_{(S,\Pi)}(T) = v_{(S_{-j}, \Pi'(j))}(T)$ if $T \subseteq S_{-j}$, and $\hat{v}_{(S,\Pi)}(T) = v_{(S_{-i}, \Pi'(i))}(T)$ if $T \subseteq S_{-i}$. $\square$

The reason why the information contained in $\hat{v}$ is sufficient to compute $\mathrm{pc}^*(\{N\})$ is more straightforward than it was in the previous sections. Indeed, applying the principle of balanced contributions (combined with partition efficiency), $\mathrm{pc}^*(\{N\})$ is entirely determined from the vectors $\mathrm{pc}_{N \setminus \{i\}}^*(\{N \setminus \{i\}, \{i\}\})$, where i varies in N. In turn, for each $i \in N$, the vector $\mathrm{pc}_{N \setminus \{i\}}^*(\{N \setminus \{i\}, \{i\}\})$ is entirely determined from the vectors $\mathrm{pc}_{N \setminus \{i,j\}}^*(\{N \setminus \{i,j\}, \{i\}, \{j\}\})$, where j varies in $N \setminus \{i\}$. Continuing in this way, only partitions with at most one atom with more than two members will matter. Nevertheless, if it is not assumed a priori that the grand coalition forms (as is the case in this section), then all the information present in the partition function could become relevant. We develop this point in the next section when we tackle coalition formation issues.

Our solution in this section is close to the logic of Aumann and Dreze (1974), who define the value for characteristic functions (no externalities) and a fixed coalition structure. In their work, as in our solution, most of the existing coalition structure is used in the calculation of the marginal contributions. However, closer to the ideas in Hart and Kurz (1983), the balanced contribution principle could also be a fruitful approach, even if the players have the possibility to join other coalitions after leaving. For instance, members of a coalition could compare their best threats against each other:

$$\sigma_i(\Pi) - \min_{T\in\Pi\setminus\{S\}} \sigma_i(\{S_{-j}, T_{+j}\} \cup \Pi_{-S,-T}) \\ = \sigma_j(\Pi) - \min_{T\in\Pi\setminus\{S\}} \sigma_j(\{S_{-i}, T_{+i}\} \cup \Pi_{-S,-T}),$$

for each embedded coalition (S, Π) and each pair (i, j) of players that belong to S. This embodies players' most pessimistic expectations of coalition formation when they leave a coalition. An inductive argument similar to the one in the last proof implies that there exists a unique payoff configuration that is partition efficient, and that satisfies this modified version of the balanced contribution principle. Of course, the same can be said if other expectations (e.g., the most optimistic) are used.

9 Coalition Formation

Now we shall explore the issue of coalition formation on the basis of our analysis in the previous section. In doing so, we combine two ways of thinking. Players may strategize in forming coalitions and predicting the coalition structure. When a coalition structure forms, however, the normative principles we have proposed are in place to sustain payoff distributions related to the value.

Of course, there are several modelling choices that one could make at this point, even if one is guided by the principle of balanced contributions that underlies the concept of the value. First, one could measure the players' contributions in different ways in order to define the appropriate payoff configuration (cf. the discussion at the end of the previous section). And second, one could model the coalition formation process itself in a variety of ways, ranging from the use of core-like ideas to the analysis of a specific extensive form, or, since the principle of balanced contributions typically leads to a unique payoff configuration, the literature on coalition formation in hedonic games might also be applied to determine the coalition structure that will emerge (see, for instance, Bogomolnaia and Jackson (2002), and references therein). Taking any of these alternative routes may lead to interesting conclusions as well.

To fix ideas here, we shall adopt the assumption used for the payoff configuration pc* identified in Proposition 4. That is, we assume that player i's contribution to player j in a coalition S equals the loss incurred by j, should i leave S. When a player leaves a coalition, he stays as a singleton in the new coalition structure. As we shall see, this assumption will not be as restrictive as it was when we imposed that the grand coalition had to form. Indeed, now all the information in the partition function will be of use in determining the answer.

In addition, we shall postulate the following extensive form to model the coalition formation process. Fix an order π for the players in N. Following the order π, each player $i \in N$ announces a coalition S such that $i \in S$. The outcome of this sequential move game of perfect information is a coalition structure, in which a coalition S forms if and only if each player in S has announced the coalition S. If Π is the coalition structure so formed, then the payoffs to players are given by $\mathrm{pc}^*(\Pi)$. Denote this extensive form game by $\Gamma(\pi)$.

Note that we could replace the payoffs that are part of pc^* with an extensive form that implements the Shapley value for each coalition in each coalition structure. For our purposes here, we prefer to take a simpler reduced form, as described. Our next result follows.

Proposition 5 *Let π be an order of the players. Then $\Gamma(\pi)$ admits at least one subgame perfect equilibrium and, for almost every game in partition function form, there exists a unique coalition structure that can be supported by a subgame perfect equilibrium of $\Gamma(\pi)$.*

Proof: This simply follows from the fact that, for almost every game v in partition function form, the payoff configuration pc^* allows no ties for each player. Thus, for almost every partition function v, the proposed coalition formation game is a finite horizon extensive form of perfect information where payoffs for each player are different at each final outcome. In such games, there exists a unique subgame perfect equilibrium outcome, which is obtained with the backwards induction algorithm. $\square$

Next, we analyze previous examples with this extensive form game, to answer the question of which coalition structure will emerge in each case.

Example 6 *Consider again the game in partition function form of Example 4:*

$$v(N) = 24;$$

$$v(\{1,2\}) = 12;$$
$$v(\{1,3\}) = 13;$$
$$v(\{2,3\}) = 14;$$
$$v(\{i\}, \{\{i\}, \{j,k\}\}) = 9 \quad \textit{for all } i, j, k;$$
$$v(\{i\}, \{\{i\}, \{j\}, \{k\}\}) = 0 \quad \textit{for all } i, j, k.$$

For this game, the payoff configuration $pc^* = (pc_1^*(\Pi), pc_2^*(\Pi), pc_3^*(\Pi))_\Pi$ of Proposition 4 yields the following:

$$pc^*(\{N\}) = (7.5, 8, 8.5);$$

$$pc^*(\{1\}, \{2,3\}) = (9, 7, 7);$$

$$pc^*(\{2\},\{1,3\}) = (6.5, 9, 6.5);$$
$$pc^*(\{3\},\{1,2\}) = (6, 6, 9);$$
$$pc^*(\{1\},\{2\},\{3\}) = (0, 0, 0).$$

Note that, since the game is superadditive, the efficient coalition structure is the grand coalition. With this payoff configuration in mind, one can now analyze our game of coalition formation.

The reader can check that, whatever the order π, the coalition structure predicted by the unique subgame perfect equilibrium has the first mover alone in his singleton coalition, anticipating that after he chooses to be alone, the other two will join together in the two-player coalition, as will surely happen. Thus, even in superadditive games the equilibrium coalition structure need not be efficient; see Ray and Vohra (1999, 2001) and Maskin (2003) for related points. Here, note how the "Coase theorem logic" does not work to solve inefficiencies by means of bargaining: the application of fairness criteria within each coalition ties the players' hands while negotiating. Of course, the specific coalition structure that emerges depends on the order, which assigns different bargaining power to players as a function of how early they speak in the game.

Example 7 *Consider now the game analyzed in Example 5. This was a variant of the game just revisited, but in which neither player 2 nor player 3 obtains any benefit from free riding. That is, the partition function is as the one just described in the previous example, but where*

$$v(\{2\},\{\{2\},\{1,3\}\}) = v(\{3\},\{\{3\},\{1,2\}\}) = 0.$$

For this game, the payoff configuration pc^* prescribes the following:

$$pc^*(\{N\}) = (7.5, 8, 8.5);$$
$$pc^*(\{1\},\{2,3\}) = (9, 7, 7);$$
$$pc^*(\{2\},\{1,3\}) = (6.5, 0, 6.5);$$
$$pc^*(\{3\},\{1,2\}) = (6, 6, 0);$$
$$pc^*(\{1\},\{2\},\{3\}) = (0, 0, 0).$$

Now one can see that, while the equilibrium coalition structure is $(\{1\},\{2,3\})$ if player 1 is the first in the order π, the grand coalition $(\{N\})$ results otherwise. The allocation of bargaining power in the order π, along with the payoff configuration pc^*, determines whether the outcome will be efficient or subject to the free rider problem.

The next example shows that superadditivity does not guarantee efficiency, even for games with no externalities.

Example 8 *Let $N = \{1, 2, 3\}$ be the set of agents and let v be the following game in characteristic function form:*

$$v(N) = 18;$$

$$v(\{1,2\}) = 16;$$
$$v(\{1,3\}) = 14;$$
$$v(\{2,3\}) = 12;$$
$$v(\{i\}) = 0 \quad \textit{for all } i.$$

Proposition 4 yields the following payoff configuration:

$$pc^*(\{N\}) = (7, 6, 5);$$

$$pc^*(\{1\}, \{2, 3\}) = (0, 6, 6);$$
$$pc^*(\{2\}, \{1, 3\}) = (7, 0, 7);$$
$$pc^*(\{3\}, \{1, 2\}) = (8, 8, 0);$$
$$pc^*(\{1\}, \{2\}, \{3\}) = (0, 0, 0).$$

The unique subgame perfect equilibrium of our game of coalition formation predicts that players 1 and 2 will cooperate, excluding player 3, whatever the order π. Any reasonable theory of coalition formation should predict the same outcome, as players 1 and 2 both rank $\{1, 2\}$ as their best alternative. In particular, the interested reader can check that $\{\{1, 2\}, \{3\}\}$ is the only core coalition structure (restricted by pc^*). Yet, its associated outcome is inefficient, as two additional dollars are available if the grand coalition forms.

A game v in characteristic function form is *strictly convex* if the marginal contributions of the players are strictly increasing in the size of the coalition with which they cooperate, i.e.,

$$v(S) - v(S_{-i}) < v(T) - v(T_{-i})$$

for each player i and each pair (S, T) of coalitions such that $i \in S$ and $S \subsetneq T$. If some of these inequalities are turned into equalities, the game in characteristic function form is said to be *convex*.

Then, we can state the following result.

Proposition 6 *If a game in characteristic function form is strictly convex, then the grand coalition forms in the unique subgame perfect equilibrium of $\Gamma(\pi)$, for each order π of the players.*

Proof: Let v be a game in characteristic function form that is strictly convex. We show that every player strictly prefers the grand coalition over any other partition. This implies that there exists a unique subgame perfect equilibrium outcome where the grand coalition forms.[3]

Let $i \in N$, let Π be a partition other than the grand coalition structure, and let S be the atom of Π to which i belongs. Proposition 4 implies that

$$\mathrm{pc}_i^*(\Pi) = \sum_{T \subseteq S \text{ s.t. } i \in T} \frac{(t-1)!(s-t)!}{s!}[v(T) - v(T_{-i})].$$

Let $T \subseteq S$ be such that $i \in T$. We have that

$$\sum_{R \subseteq N \text{ s.t. } R \cap S = T} \frac{(r-1)!(n-r)!}{n!} = \frac{(t-1)!(s-t)!}{s!}.$$

The equality follows from combinatorial calculus. Indeed, if we order all the players in N at random according to a uniform probability distribution, then the number on the left hand side of the equality represents the probability of having an order of N where a member j of S comes before i if and only if $j \in T$. If we order all the players in S at random according to a uniform probability distribution, then the number on the right hand side of the equality represents the probability of having an order of S where a member j of S comes before i if and only if $j \in T$. These two probabilities must clearly be equal.

On the other hand,

$$\begin{aligned}\textstyle\sum_{R \subseteq N \text{ s.t. } R \cap S = T} \frac{(r-1)!(n-r)!}{n!}[v(T) - v(T_{-i})] \\ < \textstyle\sum_{R \subseteq N \text{ s.t. } R \cap S = T} \frac{(r-1)!(n-r)!}{n!}[v(R) - v(R_{-i})],\end{aligned}$$

since v is strictly convex.

But notice that the left hand side is a term in $\mathrm{pc}_i^*(\Pi)$, while the right hand side is a term in $\mathrm{Sh}_i(v)$. Hence, it follows that $\mathrm{pc}_i^*(\Pi) < \mathrm{Sh}_i(v)$. $\square$

We remark that for games in characteristic function form that are convex, but not strictly convex, one can have equilibria with partitions other than the grand coalition forming. For example, suppose $v(S) = s$ for every $S \subseteq N$. For this characteristic function, every partition is supported by a subgame perfect equilibrium of the coalition formation game. It would be interesting to derive additional properties of the equilibrium coalition structure of our game $\Gamma(\pi)$.

To the best of our knowledge, Maskin (2003) is the only other paper that studies the question of coalition formation in games with externalities and whose payoff predictions are somehow related to the Shapley value. The approach is quite different, as Maskin's non cooperative procedures simultaneously determine the coalitions that form and the payoffs of the players, while we opted for

[3] Again, any reasonable theory of coalition formation should predict that the grand coalition forms, as it is unanimously considered to be the best coalition structure. In particular, it is the only core coalition structure.

a two-stage game where coalitions are formed first, and payoffs are then determined by principles of equity. The main qualitative result is similar though: inefficient outcomes may emerge in superadditive games if we introduce considerations of coalition formation. The advantage of our model is its sharp predictions and the simplicity to compute the equilibrium coalition structures in examples.

10 Conclusion

This paper has explored games in partition function form. Our basic approach is rooted in the concept of marginal contributions of players to coalitions. In games with externalities, we have argued how it is important to separate the concept of intrinsic marginal contributions from that pertaining to the externalities themselves.

The paper is divided into two parts, each corresponding to a different assumption made on the coalition structure formed. The first part of the paper assumes that the grand coalition is together. Under this assumption, the implications of anonymity, monotonicity and marginality are explored, leading to two main results. The first one establishes bounds to players' payoffs if they are to be derived from solutions that are monotonic with respect to the (weak version of) marginal contributions. The second result provides a sharp characterization of a solution that captures value-like principles, if one abstracts from the externalities. The combination of both results provides insights to the size of the Pigouvian-like transfers compatible with our normative principles.

In the second part of the paper, a solution is proposed for any coalition structure. This is achieved as a consequence of the principle of balanced contributions. Using this, a game of coalition formation is analyzed leading to a sharp prediction in most games. Our way of modelling coalition formation combines players' strategizing at the time of forming coalitions with the normative principles behind the value once a coalition structure is in place.

References

Aumann, R. J. and J. Dreze (1974), "Cooperative Games with Coalition Structures," *International Journal of Game Theory* **3**, 217-237.

Bloch, F. (1996), "Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division," *Games and Economic Behavior* **14**, 90-123.

Bloch, F. (2002), "Coalitions and Networks in Industrial Organization," *Manchester School* **70**, 36-55.

Bogomolnaia, A. and M. O. Jackson (2002), "The Stability of Hedonic Coalition Structures," *Games and Economic Behavior* **38**, 201-230.

Bolger, E. M. (1989), "A Set of Axioms for a Value for Partition Function Games," *International Journal of Game Theory* **18**, 37-44.

Fujinaka, Y. (2004), “On the Marginality Principle in Partition Function Form Games,” Mimeo, Graduate School of Economics, Kobe University, Japan.

Hart, S. and M. Kurz (1983), “Endogenous Formation of Coalitions,” *Econometrica* **51**, 1047-1064.

Hart, S. and A. Mas-Colell (1989), “Potencial, Value and Consistency,” *Econometrica* **57**, 589-614.

Hart, S. and A. Mas-Colell (1996), “Bargaining and Value,” *Econometrica* **64**, 357-380.

Macho-Stadler, I., D. Perez-Castrillo and D. Wettstein (2004), “Sharing the Surplus: An Extension of the Shapley Value for Environments with Externalities,” CREA WP 119, Barcelona, Spain.

Macho-Stadler, I., D. Perez-Castrillo and D. Wettstein (2005), “Efficient Bidding with Externalities,” CREA WP 159, Barcelona, Spain.

Maskin, E. (2003), “Bargaining, Coalitions and Externalities,” Presidential Address to the Econometric Society, Institute for Advanced Study, Princeton.

Myerson, R. B. (1977), “Value of Games in Partition Function Form,” *International Journal of Game Theory* **6**, 23-31.

Myerson, R. B. (1980), “Conference Structures and Fair Allocation Rules,” *International Journal of Game Theory* **9**, 169-182.

Perez-Castrillo, D. and D. Wettstein (2001), “Bidding for the Surplus: A Non-Cooperative Approach to the Shapley Value,” *Journal of Economic Theory* **100**, 274-294.

Pham Do, K. H. and H. Norde (2002), “The Shapley value for partition function form games,” CentER DP 202-04, Tilburg University, The Netherlands.

Ray, D. and R. Vohra (1999), “A Theory of Endogenous Coalition Structures,” *Games and Economic Behavior* **26**, 286-336.

Ray, D. and R. Vohra (2001), “Coalitional Power and Public Goods,” *Journal of Political Economy* **109**, 1355-1384.

Shapley, L. S. (1953), “A Value for n-Person Games,” in *Contributions to the Theory of Games II,* A. W. Tucker and R. D. Luce (ed.), Princeton University Press, 307-317.

Thrall, R. M. and W. F. Lucas (1963), “n-Person Games in Partition Function Form,” *Naval Research Logistics Quarterly* **10**, 281-298.

Young, H. P. (1985), “Monotonic Solutions of Cooperative Games,” *International Journal of Game Theory* **14**, 65-72.