HUB RESEARCH PAPER
Economics & Management

A New Class of Welfare Maximizing Stable Sharing Rules for Partition Function Games with Externalities

Johan Eyckmans, Michael Finus, Lina Mallozzi

HUB RESEARCH PAPER 2011/08
JULI 2011

partners in
Hogeschool-Universiteit Brussel

A New Class of Welfare Maximizing Stable Sharing Rules for Partition Function Games with Externalities*

Johan Eyckmans†
Hogeschool-Universiteit Brussel & Katholieke Universiteit Leuven, Belgium

Michael Finus
University of Exeter Business School,UK

Lina Mallozzi
University of Naples Federico II, Italy

July 15, 2011

Abstract

We propose a class of sharing rules for the distribution of the gains from cooperation for partition function games with externalities. We show that this class of sharing rules is characterized by three axioms: coalitional efficiency, additivity and anonimity which are adapted to the context of partition function games. The sharing rules stabilize, in the sense of d'Aspremont et al. (1983), the coalition which generates the highest global welfare among the set of potentially internally stable coalitions. The new class of sharing rules is particularly powerful for economic problems that are characterized by positive externalities from coalition formation (outsiders benefit from the expansion of the coalition) and which therefore often suffer from free-riding.

keywords: partition function, coalition formation, externalities, surplus sharing rules
JEL codes: C70, C71

*Valuable comments by Claude d'Aspremont, Alfred Endres, Pedro Pintasilgo and Santiago Rubio on earliers drafts of this paper are gratefully acknowledged. The usual disclaimer applies.

†Corresponding author: Hogeschool-Universiteit Brussel HUB, Center for Economic Research on Markets and their Environments, CREME, Stormstraat 2, B-1000 Brussels, Belgium, e-mail: Johan.Eyckmans@HUBrussel.be

1 Introduction

It is well-known that the presence of externalities warrants some form of coordination and cooperation between economic agents to prevent important welfare losses. By coordinating their strategies, agents can improve, both collectively and individually, upon non-cooperative outcomes provided the gains from cooperation are shared in accordance with the individual interests of all participants. Examples of economic problems with externalities are plentiful in the literature and range from output and price setting behavior of firms in imperfectly competitive markets (e.g. Deneckere and Davidson 1985 and Yi 1997), R&D investments of firms (e.g. Poyago-Theotokay 1995 and Yi and Shin 2000), the provision of public goods, like greenhouse gas emission reduction, the eradication of contagious diseases, like Malaria and AIDS, and the fight against terrorism (e.g. Sandler 2004 for an overview), to coordination of tariffs and monetary policies (e.g. Yi 1996 and Kohler 2002).

The classical approach of studying the formation of coalitions assumes a transferable utility (TU)-framework and is based on the characteristic function. This function assigns to every coalition a worth which is the aggregate payoff that a coalition can secure for its members, irrespective of the behavior of players outside this coalition. By construction, the characteristic function approach is ill-suited to capture externality problems, *i.e.* outsiders' actions affect the worth of a coalition and vice versa (Bloch 2003). The focus is on the coalition's worth, and how it might be shared in the grand coalition, but nothing is said about the payoff to outsiders. An alternative approach is based on the concept of the partition function introduced by Thrall and Lucas (1963). This function assigns a worth to every coalition and singleton in the game. The worth depends on the entire coalition structure, *i.e.* the partition of players. The main focus of this literature is on the prediction of equilibrium coalition structures emerging from some coalition formation process. Since in our analysis externalities are a central feature, we work in the spirit of the partition function approach, considering explicitly the payoffs of all players, regardless whether they are involved in a coalition.

Beyond the particularities of the underlying economic models, it appears that games in partition function form share some common features in terms of the incentives to form coalitions. As shown for instance by Yi (1997), Bloch (2003) and Maskin (2003), a crucial feature for the success of coalition formation is the sign of the externality that accrues to those players not involved in the expansion of a coalition. It appears that positive externalities from coalition formation provide an incentive for players to free-ride, leading to only small stable coalitions. Typical examples of positive externalities include for instance output and price cartels and the provision of public goods. Firms not involved in an output cartel benefit from lower output by the cartel via higher market prices. This is also the driving force in price cartels where the cartel raises prices above non-cooperative levels. Similarly, agents not involved in a public good agreement benefit from higher provision levels of participants. Given the fact that

cooperation is most difficult to achieve in positive externality games, we will focus on this class of games, although almost all of our results also carry over to negative externality games.[1]

Intuitively, it is clear that the degree and success of cooperation depends crucially on the division of the gains from cooperation. There is a substantial body of literature on the axiomatic underpinning of surplus sharing rules in the context of the partition function approach; see for instance Meyerson (1977), Bolger (1989), de Clippel and Serrano (2008), Macho-Stadler, Pérez-Castrillo and Wettstein (2007) and Grabisch and Funaki (2008). However, this literature often assumes cooperation among all agents (*i.e.* grand coalition) to be stable per se, despite the fact that there are plenty of economic problems where it seems reasonable to expect that the grand coalition will not be stable because of free-riding behavior (e.g. Ray and Vohra 1999, Maskin 2003 and Hafalir 2007). In contrast, we consider stability as a necessary condition for sustainable cooperation. Coalitional stability is an equilibrium outcome of our sharing scheme, not an a priori assumption.

Up to now, the literature using the partition function has paid little attention to the division of the gains from cooperation. Due to the complexity of partition functions, most papers assume a fixed sharing rule. One set of papers, which comprises the bulk of the literature, assumes ex-ante symmetric players with equal sharing (e.g. Bloch 2003, Yi and Shin 2000 and Ray and Vohra 2001). This is mainly done for analytical tractability. However, symmetry is a strong assumption that is difficult to justify in most economic environments. Another set of papers — mainly related to the analysis of public good agreements — allows for asymmetric players but makes ad hoc assumptions about particular (exogenous) sharing rules, most of which are classical solution concepts or modifications of them (e.g. Nash Bargaining Solution or Shapley Value; Barrett 1997 and Botteon Carraro 1997). Clearly, this approach is also not satisfactory for at least three reasons. First, none of these sharing rules guarantees existence of an equilibrium coalition structure. Second, the prediction of equilibrium coalition structures is sensitive to the particular specification of the sharing rules. Third, it remains an open question whether there are other sharing rules that could perform better in terms of the aggregate worth, letting alone whether there is a sharing rule that is "optimal". Optimal means to achieve the maximum aggregate worth subject to the condition that the underlying coalition structure is stable. In order to address these problems, we propose a new class of sharing rules, the Proportional Surplus Sharing Scheme, abbreviated PSSS, and have a closer look at one of its members, the Equal Surplus Sharing Scheme (ESSS).

Our analysis is based on the well-known concept of internal and external stability introduced by d'Aspremont et al. (1983) in the context of cartel formation. Players are assumed to have only the choice between remaining at the fringe (*i.e.* forming

[1]Examples of negative externality games are discussed in Yi (1997) and Bloch (2003).

a singleton coalition) or joining the cartel (*i.e.* forming a non-trivial coalition). The cartel is called stable if no cartel member has an incentive to leave (internal stability) and no outsider has an incentive to join (external stability) the cartel. On the one hand, focussing on a single non-trivial coalition is a restriction. On the other hand, however, it allows deriving conclusive results under minimal structural assumptions: no assumption about the heterogeneity of players is required and only very general and mild structural properties of the partition function are necessary. Moreover, the concept of internal and external stability has the advantage that the set of stable coalitions can be interpreted as the set of Nash equilibria of a simple Yes-No announcement game of coalition formation.

A central concept in our analysis is the notion of Potentially Internally Stable (PIS) coalition structures. Loosely speaking, PIS coalition structures generate sufficient surplus for coalition members in order to satisfy their free-rider claims. In the presence of externalities, this may only be possible for a subset of all possible coalition structures. Hence, our new sharing scheme does not necessarily lead to efficiency in the sense that the grand coalition forms. This is particularly relevant for positive externality games in which large coalitions are typically not PIS due to strong free-rider incentives.

It will become apparent that the structure of our sharing scheme resembles the Nash Bargaining solution but where threat points depend on the coalition structure.[2] Every coalition members receives his threat point payoff plus a share of the surplus which is the difference between the aggregate worth and the sum of free-rider payoffs. The shares are represented by weights with the PSSS allowing for any positive set of weights and ESSS focusing on the case of equal weights.

For the PSSS we show that a stable (internally and externally) coalition always exists and, most importantly, among the set of PIS coalition structures, the coalition structure with the highest aggregate worth is stable. For the ESSS we demonstrate that it is characterized by axioms in the same spirit of those typically used to characterize the Shapley value and its extensions in partition function games. The ESSS is characterized by versions of coalitional efficiency, additivity and anonimity axioms. In that sense, the ESSS is close to the axiomatic surplus sharing literature, though without assuming that the grand coalition will always form. In other words, the ESSS bridges two strands of literature: the cooperative and axiomatic surplus sharing literature (e.g. Meyerson 1977, Bolger 1989, de Clippel and Serrano 2008, Macho-Stadler, Pérez-Castrillo and Wettstein 2007) and the non-cooperative coalition formation literature (e.g. Bloch 2003, Ray and Vohra 2001, Maskin 2003 and Faigle and Grabish 2011).

In the following, we first introduce notations and definitions in section 2. Section 3 introduces our new sharing scheme PSSS and proves its properties. The axiomatization

[2] A similar idea in the context of international environmental agreements has recently been advanced by Fuentes-Albero and Rubio (2010), McGuinty (2007) and Weikard (2007).

of this sharing scheme for equal weights, ESSS, is the subject of section 4. Finally, section 5 summarizes the main findings and points to some issues for future research.

2 Ingredients

2.1 Partition Function

We denote by $N = \{1, ..., n\}$ $(n \geq 2)$ the set of players in the game. We consider coalition structures, *i.e.* partitions of the player set N, consisting of only one non-empty coalition $S \subseteq N$ with $\#S = s > 0$ (the cartel or coalition) while all other players $j \in N \setminus S$ are singletons (the fringe). Let $\mathcal{B}$ be the set of all possible coalition structures of this type $\mathcal{B} = \{\beta_S = (S, \{j_1\}, ..., \{j_{n-s}\}),\ S \subseteq N,\ j_1, ..., j_{n-s} \in N \setminus S\}$. We define a partition function π that assigns a single real number $\pi_S(\beta_S)$ to coalition S and real numbers $\pi_{\{j\}}(\beta_S)$ to every singleton $j \in N \setminus S$ of the fringe as follows:

$$\pi : \beta_S \in \mathcal{B} \mapsto \pi(\beta_S) = (\pi_S(\beta_S), \pi_{\{j_1\}}(\beta_S), ..., \pi_{\{j_{n-s}\}}(\beta_S))$$

For notational simplicity, for any $j \in N$, we denote $\pi_{\{j\}}(\beta_S)$ by $\pi_j(\beta_S)$ in the sequel. The image of this mapping is a vector of variable size $1 + (n - s)$, depending on the cardinality of coalition S and on the total number of players. On the one hand, our partition function is simpler than general partition functions (see, e.g., Bloch 2003 and Yi 2003) since we disregard all partitions that consist of two or more non-trivial coalitions. On the other hand, and in contrast to the classic characteristic function, our partition function assigns not only a worth to coalition S but also to the non-members of S. This is important because information on the payoffs to players outside the coalition is indispensable for analyzing the incentives to leave or join a coalition in games with externalities.

We define a transferable utility partition function game $\Gamma = (N, \mathcal{B}, \pi)$, consisting of a player set $N = \{1, ..., n\}$ (with $n \geq 2$), a set of partitions or coalition structures $\mathcal{B}$ and a partition function π mapping to each coalition structure $\beta_S \in \mathcal{B}$ a real valued vector representing the worth of all the coalitions in β_S. Following, among others, Yi and Shin (2000), Maskin (2003) and Hafalir (2007), we introduce a property of the partition function that will prove particularly useful in establishing results later for different sub-classes of partition function games.

Definition 1 : **Positive and Negative Externalities**

A game in partition function form $\Gamma(N, \mathcal{B}, \pi)$ *exhibits positive (negative) externalities if and only if its partition function* π *satisfies:* $\forall S \subseteq N, \forall j \in N \setminus S, j \neq i : \pi_j(\beta_S) \geq (\leq)\pi_j(\beta_{S\setminus\{i\}})$ *and* $\exists k \in N \setminus S, k \neq i : \pi_k(\beta_S) > (<)\pi_k(\beta_{S\setminus\{i\}})$.

That is, a positive externality implies that no outsider, *i.e.* a player that is not involved in the enlargement of the coalition, is worse off and at least one outsider is strictly better off. The reverse holds for a negative externality. Note that a game is said to be a positive (negative) externality game if this property holds for all possible coalition structures in $\beta_S \in \mathcal{B}$. The general underlying structure of externality games is that the payoff function of at least one player does not only depend on his own economic strategies (e.g. output, prices, tariffs and R&D investment) but also on at least one other player's strategy. Consequently, forming the grand coalition and maximizing the aggregate worth, is the natural benchmark against which the success of coalition formation can be measured.

2.2 Valuation Function

For the analysis of the incentives of individual players to form coalitions, we have to take one more step: it is not the aggregate payoff to a coalition but the individual payoffs to coalition members that matter. This type of information is contained in the *valuation function*, *i.e.* a function mapping coalition structures into a vector of individual payoffs, called valuations. We define a valuation function that assigns to every coalition structure $\beta_S \in \mathcal{B}$ a real-valued vector $f^\pi(\beta_S) \in \mathbb{R}^n$ as follows:

$$\begin{cases} \sum_{i \in S} f_i^\pi(\beta_S) = \pi_S(\beta_S) & , \\ f_j^\pi(\beta_S) = \pi_j(\beta_S) & \text{if } j \notin S. \end{cases}$$

This means that for every coalition S, the valuation function f^π specifies how the worth of coalition S is allocated among its members. By construction, valuations are group rational, *i.e.* the entire worth $\pi_S(\beta_S)$ of coalition S is distributed among its members. For every outsider to coalition S, the valuation coincides with the worth $\pi_j(\beta_S)$ that is assigned to him by the partition function.

Obviously, there are many possibilities to construct valuation functions starting from a particular partition function. In the sequel, we will propose a particular family of valuation functions with desirable properties for games with externalities.

Note the difference between the concept of a valuation and an imputation known from the characteristic function. An imputation is usually only one vector of length n, listing the payoff of each player in the grand coalition ($S = N$) whereas a valuation function assigns vectors of length n to every possible coalition structure, specifying individual payoffs not only for coalition members but also for outsiders.[3] This more comprehensive view is necessary to capture externalities across coalitions and players.

[3] An exception is Aumann and Drèze (1974) who consider various solution concepts not only for the grand coalition but also for any partition of players. We will apply one of their concepts to the example we introduce below.

Given a valuation function containing information on individual players' payoffs, we can now introduce the notion of stable coalitions following d'Aspremont et al. (1983).

Definition 2 : **Internal and External Stability**

Let f^π be a valuation function for the game in partition function form $\Gamma(N, \mathcal{B}, \pi)$ and $f(\beta_S)$ the vector of valuations for the players in N when coalition structure β_S forms. Coalition structure $\beta_S \in \mathcal{B}$ is stable *with respect to the valuation function f^π if and only if:*

$$\textit{internal stability:} \quad \forall i \in S : f_i^\pi(\beta_S) \geq f_i^\pi(\beta_{S\setminus\{i\}}),$$

$$\textit{external stability:} \quad \forall j \in N \setminus S : f_j^\pi(\beta_S) \geq f_j^\pi(\beta_{S\cup\{j\}}).$$

That is, coalition structure β_S is stable if it is internally (IS) and externally (ES) stable, *i.e.* no insider wants to leave and no outsider wants to join coalition S.

As there are many ways to share the coalitional worth in a coalition game Γ, there are many possible valuation functions that can be derived from its partition function. Consequently, a partition β_S may be stable with respect to a particular valuation function f^π but may not be stable with respect to another valuation function g^π. Therefore, we denote the set of coalition structures that are internally stable with respect to valuation function f^π by $\Sigma^{IS}(f^\pi)$, the set of coalition structures that are externally stable by $\Sigma^{ES}(f^\pi)$ and the set of stable coalition structures by $\Sigma^S(f^\pi) = \Sigma^{IS}(f^\pi) \cap \Sigma^{ES}(f^\pi)$.

In order to facilitate the subsequent analysis, it will prove useful to classify coalition structures according to the following criterion.

Definition 3 : **Potential Internal Stability**

Consider a game in partition function form $\Gamma = (N, \mathcal{B}, \pi)$. A coalition structure $\beta_S \in \mathcal{B}$ is called Potentially Internally Stable for game Γ if and only if:

$$\pi_S(\beta_S) \geq \sum_{i\in S} \pi_i(\beta_{S\setminus\{i\}}).$$

That is, a coalition structure is Potentially Internally Stable (PIS) if the aggregate payoff to the coalition is not smaller than the sum of free-rider payoffs of its members. The free-riding payoff is the payoff that a player can achieve if he leaves coalition S and the remaining members continue cooperating in $S \setminus \{i\}$. Note that the property of potential internal stability refers to the partition function as only aggregate payoffs matter whereas internal stability and external stability are (and have to be) properties of a valuation function as individual payoffs matter. For later reference, we will denote the set of coalitions that are PIS for a particular partition function by π by $\Sigma^{PIS}(\pi)$.

3 Proportional Surplus Sharing Scheme

3.1 Motivation

In this section, we introduce our new surplus sharing scheme for partition function games with externalities, the Proportional Surplus Sharing Scheme (PSSS). Since it is our objective to study the impact of surplus sharing rules on the stability of coalitions in a general framework, the use of one specific sharing rule would be too restrictive. Therefore, we introduce a class of sharing rules and study the properties of its members. Recall from the definition of a valuation function that there corresponds to every surplus sharing rule a particular valuation function. Hence, speaking about a sharing rule is equivalent to speaking about a valuation function and it is the latter terminology that we will use in the remainder of this paper.

The main motivation for the development of the PSSS is to remedy free-riding in the context of positive externalities. In the context of positive externalities, associated with the problem of forming large stable coalitions, it appears that one is more concerned about players leaving the coalition than about players joining it. Hence, the most immediate notion of free-riding seems to be related to the violation of internal stability, although we have to be aware that stability also comprises external stability. Therefore, our analysis starts from internal stability and subsequently we consider also external stability.

In order to illustrate some results in the course of the following discussion, we provide a numerical example in the Appendix. The example assumes four players and a partition function that exhibits positive externalities, see Table A.1 in the Appendix. As can be seen from column (10), all coalition structures comprising two-player coalitions, and two out of four coalition structures comprising a three-player coalition are PIS. The other two three-player coalitions and the grand coalition do not generate enough surplus to compensate for the free-riding claims of their members. The challenge for a surplus sharing rule is to achieve maximal global welfare and to ensure stability of the coalition structure at the same time.

Meeting this challenge is far from obvious as this is illustrated in Table A.2. Here valuations for the Shapley Value are reported as defined in Aumann and Drèze (1974), *i.e.* these values are computed not only for the grand coalition but for every non-empty coalition. Two interesting points can be observed in this example. First, a stable coalition may not always exist. Second, there is no indication as to whether there are other sharing rules leading to stable coalitions and in particular those with higher aggregate worth. In the following, we claim that there is such a sharing rule: the PSSS.

3.2 Definitions

The construction of the PSSS starts from the observation that a necessary and sufficient condition for internal stability of coalition structure β_S is that each player in S receives his outside payoff. In the cartel formation game, the free-rider payoffs are associated with the scenario in which an individual coalition member leaves coalition S in order to become a singleton while the remaining members of coalition S continue to cooperate. These payoffs constitute lower bounds on the claims of individual coalition members with respect to the coalitional surplus in order to refrain from leaving the coalition.

Definition 4 : **Proportional Surplus Sharing Valuation Function**

A Proportional Surplus Valuation Function for a game in partition function form $\Gamma = (N, \mathcal{B}, \pi)$ *is a valuation function* v^π *that satisfies* $\forall S \subseteq N$*:*

$$\begin{cases} v_i^\pi(\beta_S) = \pi_i(\beta_{S\setminus\{i\}}) + \lambda_i(\beta_S)\sigma(\beta_S) & \forall i \in S \\ v_j^\pi(\beta_S) = \pi_j(\beta_S) & \forall j \in N \setminus S \end{cases}$$

with $\lambda(\beta_S) \in \Delta^{s-1} = \{\lambda \in \mathbb{R}_+^s \mid \sum_{i\in S} \lambda_i = 1\}$ *and* $\sigma(\beta_S) = \pi_S(\beta_S) - \sum_{i\in S} \pi_i(\beta_{S\setminus\{i\}})$*.* Δ^{s-1} *denotes the set of all possible positive sharing weights of a coalition with* s *players and* $\sigma(\beta_S)$ *denotes the surplus or deficit of coalition* S *over the sum of free-rider payoffs* $\pi_i(\beta_{S\setminus\{i\}})$ *of its members.*

Intuitively, the Proportional Surplus Sharing Valuation Function (PSSVF) allocates to each coalition member his free-rider payoff, plus a non-negative share of the remaining surplus (or deficit) in proportion to weights $\lambda(\beta_S)$. Differences in valuations of players can be the result of (i) different free-rider payoffs, and/or (ii) different weights according to which the surplus or deficit is shared.

One obvious interpretation of our definition is that the free-rider payoff $\pi_i(\beta_{S\setminus\{i\}})$ is the threat point of player i in coalition S and weight $\lambda_i(\beta_S)$ is his bargaining power. Hence, the PSSVF can be seen as an extension of a Nash bargaining solution in TU games in the context of games in partition function form. However, different from the Nash bargaining solution, the threat point is not fixed but depends on the coalition structure. For instance, if S gradually increases through the accession of players, the threat point will also gradually increase (decrease) as a result of positive (negative) externalities.

It is important to point out that there are as many PSSVFs for a game in partition function form as there are ways to share — in every possible coalition S of N — the coalition surplus among its members. The set of all PSSVFs for game Γ will be denoted by $\mathcal{V}(\Gamma)$ and constitutes the PSSS. Despite that every PSSVF is defined for specific weights λ, the class of PSSVFs, constituting the PSSS, does not require assumptions

about weights as long as they are non-negative and sum up to one for each coalition S of N. This stresses that — different from most cooperative solution concepts, *e.g.*, Shapley Value and Nucleolus — the PSSS does not require to assign a unique value to each player in some coalition S. Instead, we find it more appealing that weights do not matter for many of our subsequent results, *i.e.* several results hold for the entire class of PSSVFs.

Table A.3 displays results in the numerical example for our PSSVF. The PSSVF achieves internal stability for all nine PIS coalitions (*i.e.* compare Table A.1, column 10 with Table A.3, column 7). This observation will be generalized in Lemma 1 below. In contrast, the Shapley Value can only ensure internal stability of the singleton coalition structure and of all two-player coalitions (Table A.2, column 7) and hence only of seven out of nine potentially internally stable coalitions. That is, the two PIS coalitions with three members are not internally stable for the Shapley Value.

3.3 Preliminary Properties

The construction of PSSVFs in Definition 4 suggests that it is designed to remedy free-riding in terms of internal stability and this was confirmed in the numerical example. The following Lemma 1 formalizes and generalizes this observation.

Lemma 1 : **Potential Internal Stability of PSSVFs**

Consider a partition function form game $\Gamma(N, \mathcal{B}, \pi)$. *Partition* $\beta_S \in \mathcal{B}$ *is potentially internally stable if and only if it is internally stable for any* $v^\pi \in \mathcal{V}(\Gamma)$.

Proof:
$\Longrightarrow$ Suppose that coalition structure $\beta_S \in \mathcal{B}$ is PIS, implying $\sigma(\beta_S) \geq 0$, but assume to the contrary that there exists a valuation $v^\pi \in \mathcal{V}(\Gamma)$ such that partition $\beta_S \in \mathcal{B}$ would not be internally stable with respect to this valuation function. Thus, $\pi_i(\beta_{S\setminus\{i\}}) + \lambda_i(\beta_S)\sigma(\beta_S) = v_i^\pi(\beta_S) < v_i^\pi(\beta_{S\setminus\{i\}}) = \pi_i(\beta_{S\setminus\{i\}})$ which would imply $\sigma(\beta_S) < 0$ and therefore contradicts the initial assumption.

$\Longleftarrow$ If β_S is internally stable for any valuation $v^\pi \in \mathcal{V}(\Gamma)$, then it follows that $v_i^\pi(\beta_S) \geq v_i^\pi(\beta_{S\setminus\{i\}}) = \pi_i(\beta_{S\setminus\{i\}})$. Taking sums over all members in S and using the definition of PSSVF yields $\sum_{j\in S}\left[\pi_j(\beta_{S\setminus\{j\}}) + \lambda_j(\beta_S)\sigma(\beta_S)\right] \geq \sum_{j\in S}\pi_j(\beta_{S\setminus\{j\}})$ and hence $\sigma(\beta_S) \geq 0$. From Definition 4, it therefore follows that β_S is PIS. □

The importance of Lemma 1 derives from three facts. First, internal stability is a necessary condition for a stable coalition, but is often violated for larger coalitions in positive externality games as this appears from the literature (e.g. Maskin 2003

and Hafalir 2006). Second, every PSSVF ensures that every coalition structure that is PIS will actually be internally stable. Other solution concepts may miss this potential substantially as we already observed in the numerical example. Third, there is a high degree of freedom in the choice of the sharing rule (through the choice of weights λ) when aiming at stabilizing coalitions internally.

The next Lemma 2 will turn out to be very useful in the sequel because it establishes an important link between internal and external stability for every member of the PSSS.

Lemma 2 : **External and Potential Internal Stability of PSSVFs**

Consider a game in partition function form $\Gamma(N, \mathcal{B}, \pi)$ *and a valuation function* $v^\pi \in \mathcal{V}(\Gamma)$. *If coalition structure* β_S *is not externally stable with respect to* v^π, *then there exists a player* $j \in N \setminus S$ *such that coalition structure* $\beta_{S\cup\{j\}}$ *is potentially internally stable.*

Proof:
If coalition structure β_S is not externally stable with respect to v^π, then it follows from Definition 2 that: $\exists j \in N \setminus S : v_j^\pi(\beta_{S\cup\{j\}}) > v_j^\pi(\beta_S)$. Using Definition 4 of PSSVF this is equivalent to $\pi_j(\beta_S) + \lambda_j(\beta_{S\cup\{j\}})\sigma(\beta_{S\cup\{j\}}) > \pi_j(\beta_S)$ or $\sigma(\beta_{S\cup\{j\}}) > 0$, implying that $S \cup \{j\}$ is PIS. □

It is important to note that Lemma 2 is a distinctive property of the class of PSSVFs and may not hold for other valuation functions, for instance valuation functions based on traditional cooperative solutions like the Shapley value as this is evident from the numerical example in the appendix (see Table A.2).

Using Lemma 1 and 2, we now establish our first main result, the existence of a stable coalition structure for every PSSVF.

Proposition 1 : **Existence of a Stable Coalition Structure**

Consider a game in partition function form $\Gamma(N, \mathcal{B}, \pi)$ *and the corresponding class of PSSVFs* $\mathcal{V}(\Gamma)$. *For any* $v^\pi \in \mathcal{V}(\Gamma)$ *there exists at least one stable coalition structure* $\beta_S \in \mathcal{B}$.

Proof:
By definition, the trivial coalition structure $(\{1\}, ..., \{n\})$ is internally stable. If it is also externally stable, we are done. Suppose, however, that this is not the case. Then there exists at least one two-player coalition that is PIS by Lemma 2 and which is internally stable for any $v^\pi \in \mathcal{V}(\Gamma)$ by Lemma 1. Again, if one of these two-player coalitions is also externally stable, we are done. Continuing with this reasoning, it is

evident that some coalition $S \subseteq N$ will be eventually internally and externally stable, noting that $S = N$ is externally stable by definition. □

It should be noted that the line of reasoning in this proof follows closely the argument in d'Aspremont et al. (1983). However, there is an important difference because of Lemma 2, we neither have to assume symmetric valuations for all players in S, *i.e.* $v_i^\pi(\beta_S) = v_j^\pi(\beta_S)\ \forall i, j \in S$, nor symmetric valuations for all non-member, *i.e.* $v_k^\pi(S) = v_l^\pi(S)\ \forall k, l \in N \setminus S$, as this is done in d'Aspremont et al. (1983) and in much of the literature (see Bloch 2003 and Yi 1997 for example). Also no structural assumption about the underlying economic model is required and hence also not about the properties of the partition function game.

From the proof of Proposition 1 it is evident that it would be sufficient if there exists at least one non-trivial (*i.e.* at least a two-player or possibly larger) PIS coalition, to ensure existence of a non-trivial stable coalition structure under the PSSS. In that case, we would have a starting point to apply the "algorithmic existence proof" above. In many applications such non-trivial PIS coalitions do exist. On a more general level, note that in games characterized by superadditivity[4], all two-player coalitions are PIS and can serve as a starting point for the existence proof. Hence, superadditivity is a sufficient, but not a necessary condition for existence of at least one non-trivial stable coalition structure under the PSSS.

3.4 Maximal Welfare Stable Coalition Structure

We now turn to our central result related to welfare optimality. It says that adopting a PSSS guarantees that the coalition structure which generates the highest aggregate worth (*i.e.* total payoff over all players) among all PIS coalition structures will not only be internally stable but also externally stable and therefore stable. The remarkable aspect of this result is that a sharing scheme that, at first sight, is designed to foster internal stability only, is also capable of ensuring external stability for those coalitions that are most desirable in terms of the aggregate welfare.

Proposition 2 : **Maximal Welfare Stable Coalition Structure**

Let $\Sigma^{PIS}(\pi)$ be the set of coalition structures that are potentially internally stable in partition function form game $\Gamma(N, \mathcal{B}, \pi)$ with positive externalities and let β_{S^*} be the coalition structure with the highest aggregate worth in $\Sigma^{PIS}(\pi)$: $\pi_{S^*}(\beta_{S^*}) + \sum_{j \in N \setminus S^*} \pi_j(\beta_{S^*}) \geq \pi_S(S) + \sum_{j \in N \setminus S} \pi_j(S)$. Then, every valuation function $v^\pi \in \mathcal{V}(\Gamma)$ will make coalition structure β_{S^*} both (i) internally and (ii) externally stable and, hence, stable.

[4]A game is called superadditive if the corresponding partition function satisfies $\pi_S(\beta_S) \geq \pi_{S\setminus\{i\}}(\beta_{S\setminus\{i\}}) + \pi_i(\beta_{S\setminus\{i\}})$ for all $S \subset N$ and for all $i \in S$.

Proof:
(i) Follows from Lemma 1.

(ii) Assume to the contrary that $\beta_{S^*} \in \Sigma^{PIS}(\pi)$ would not be externally stable for some valuation function $v^\pi \in \mathcal{V}(\pi)$. Hence, it follows from Definition 2 that there is an outsider $j \in N \setminus S^*$ who would strictly gain from joining coalition S^*: $v_j^\pi(\beta_{S^*\cup\{j\}}) > v_j^\pi(\beta_{S^*})$. Using Definition 4, it follows that $\pi_j(\beta_{S^*}) + \lambda_j(\beta_{S^*\cup\{j\}})\sigma(\beta_{S^*\cup\{j\}}) > \pi_j(\beta_{S^*})$ and therefore we have that $\sigma(\beta_{S^*\cup\{j\}}) > 0$ which is equivalent to $\sum_{k\in S^*\cup\{j\}}\pi_k(\beta_{S^*\cup\{j\}}) > \sum_{k\in S^*\cup\{j\}}\pi_k(\beta_{S^*})$. Hence, members of coalition $S^* \cup \{j\}$ are strictly better off under $S^* \cup \{j\}$ than under S^* and $S^* \cup \{j\}$ is PIS. At the same time, we know from the positive externalities property that outsiders to $S^* \cup \{j\}$ are not worse off under $S^* \cup \{j\}$ than under S^*: $\sum_{k\in N\setminus(S^*\cup\{j\})}\pi_k(\beta_{S^*\cup\{j\}}) \geq \sum_{k\in N\setminus(S^*\cup\{j\})}\pi_k(\beta_{S^*})$. Combining the inequalities for members and non-members of coalition S it would hold that $\sum_{k\in N}\pi_k(\beta_{S^*\cup\{j\}}) > \sum_{k\in N}\pi_k(\beta_{S^*})$, contradicting the initial assumption that S^* generates the highest worth among all coalition structures that are PIS. □

Proposition 2 can be interpreted as saying that we cannot do better in terms of the total payoff than adopting PSSS if the agreement is required to satisfy stability in the sense of d'Aspremont et al. (1983). Note first that this result is very general since it only requires that the underlying game exhibits positive externalities. There is no need for further assumptions like superadditivity and hence this result applies to many interesting economic problems. Second, since a PSSS is a collection of parametric valuation function, related to the specific set of sharing weights $\lambda(\beta_S), \forall\beta_S \in \mathcal{B}$, there remains considerable flexibility how to allocate the surplus of the coalition without jeopardizing optimality.

In the numerical example in the Appendix, the PSSVF stabilizes the three player coalition structure $(\{a, b, c\}, \{d\})$ with the highest aggregate payoff among all potentially internally stable coalitions (*i.e.* $13\frac{1}{3}$), achieving 75.5% of the maximal aggregate payoff which would be obtained by the grand coalition (*i.e.* $17\frac{2}{3}$), see table A.3. In contrast, the Shapley value (see table A.2) in this example leads to seven internally stable coalition structures of which none is also externally stable. Hence, there are no stable coalitions for the Shapley value, *i.e.* this solution concept cannot achieve sustainable cooperation.

4 Characterization

So far, we have shown existence and optimality of a stable coalition structure under PSSS. Another important question, however, is how the new sharing scheme relates to existing cooperative solutions. In order to facilitate this comparison, we provide

a characterization of the PSSS in the spirit of the Shapley value, using axioms like coalitional efficiency, anonimity and addivity. However, as the standard axioms are defined for the cooperative game framework, we have to adapt them to our context of partition function games.

4.1 Coalitional Efficiency

Axiom 1 : **Coalitional Efficiency**

Consider a game in partition function form $\Gamma(N, \mathcal{B}, \pi)$*. A valuation function* f^π *is called coalitionally efficient if and only if for all* $\beta_S \in \mathcal{B}$ *it holds that*

$$\begin{cases} \sum_{i \in S} f_i^\pi(\beta_S) = \pi_S(\beta_S) \\ f_j^\pi(\beta_S) = \pi_j(\beta_S) \qquad\qquad \text{for any } j \in N \setminus S. \end{cases}$$

Coalitional Efficiency (CE) requires that, for every possible coalition structure $\beta_S \in \mathcal{B}$, the value of the coalition S in coalition structure β_S is fully distributed among its members and that outsiders get their individual payoff. This axiom is important as it ensures that no surplus is wasted in the sharing scheme. By construction, see Definition 4, the PSSS satisfies this axiom.

Proposition 3 : **Coalitional Efficiency of PSSS**

Consider a game in partition function form $\Gamma(N, \mathcal{B}, \pi)$*. The PSSS, by definition of the PSSVF* $v^\pi \in \mathcal{V}(\Gamma)$*, satisfies the coalitional efficiency axiom for any coalition structure* $\beta_S \in \mathcal{B}$*.*

4.2 Additivity

A second frequently used axiom for a value or solution is Additivity (AD). Consider two games with partition functions π^1 and π^2. Loosely speaking, addivity requires that the outcome of the combined game characterized by the sum of the partition functions $\pi^1 + \pi^2$ is the same as the sum of the outcomes of both games separately.

Axiom 2 : **Additivity**

Consider two games in partition function form $\Gamma(N, \mathcal{B}, \pi^1)$ *and* $\Gamma(N, \mathcal{B}, \pi^2)$*. A valuation function* f^π *is called additive if and only if for any coalition structure* $\beta_S \in \mathcal{B}$ *and for every* $i \in N$*, it holds that*

$$f_i^{\pi^1+\pi^2}(\beta_S) = f_i^{\pi^1}(\beta_S) + f_i^{\pi^2}(\beta_S).$$

The additivity axiom may be motivated in several ways. For example, we may interprete the partition function as an expected payoff. Then additivity is desirable because one can add the value of different states of the world. Also in the case of cost-sharing games where the game corresponds to sharing the cost of several services, the additivity is desirable because the cost of a joint service should be the sum of the cost of separate services.[5]

Proposition 4 : Additivity of PSSS

The PSSS satisfies additivity for any pair π^1, π^2 of partition functions.

Proof:
Consider two partition functions π^1 and π^2. We define the sum of the partition functions $\bar{\pi} = \pi^1 + \pi^2$ such that for all $\beta_S \in \mathcal{B}$ it holds that (i) $\bar{\pi}_S(\beta_S) = \pi^1_S(\beta_S) + \pi^2_S(\beta_S)$ and (ii) for all $j \notin S$ $\bar{\pi}_j(\beta_S) = \pi^1_j(\beta_S) + \pi^2_j(\beta_S)$. For any $i \in S$ we have:

$$
\begin{aligned}
v_i^{\bar{\pi}}(\beta_S) &= \bar{\pi}_i(\beta_{S\setminus\{i\}}) + \lambda_i(\beta_S)\left[\bar{\pi}_S(\beta_S) - \sum_{j\in S}\bar{\pi}_j(\beta_{S\setminus\{j\}})\right] \\
&= \left[\pi^1_i(\beta_{S\setminus\{i\}}) + \pi^2_i(\beta_{S\setminus\{i\}})\right] + \\
&\qquad \lambda_i(\beta_S)\left[\left[\pi^1_S(\beta_S) + \pi^2_S(\beta_S)\right] - \sum_{j\in S}\left[\pi^1_j(\beta_{S\setminus\{j\}}) + \pi^2_j(\beta_{S\setminus\{j\}})\right]\right] \\
&= \left[\pi^1_i(\beta_{S\setminus\{i\}}) + \lambda_i(\beta_S)\left[\bar{\pi}^1_S(\beta_S) - \sum_{j\in S}\bar{\pi}^1_j(\beta_{S\setminus\{j\}})\right]\right] + \\
&\qquad \left[\pi^2_i(\beta_{S\setminus\{i\}}) + \lambda_i(\beta_S)\left[\bar{\pi}^2_S(\beta_S) - \sum_{j\in S}\bar{\pi}^2_j(\beta_{S\setminus\{j\}})\right]\right] \\
&= v_i^{\pi^1}(\beta_S) + v_i^{\pi^2}(\beta_S).
\end{aligned}
$$

□

Note that this proof hinges crucially on the assumption that the weights $\lambda_i(\beta_S)$ do not depend on the partition function π. They can, however, depend on other parameters, like for instance the number of players in the game.

4.3 Anonimity

Anonymity (AN) requires that the outcome of the sharing scheme does not depend on the identity or name of the players.

[5]Cases where additivity does not hold are discussed for instance in Kolpin (1996).

First, this axioms requires more of an adaptation than the previous two axioms, coalitional efficiency and additivity, compared to their traditional meaning in cooperative game theory. In particular, we cannot allow for any possible permutation of agents because this would upset the coalition structure of the partition. We therefore follow Aumann and Dreże (1974) and restrict permutations to respect to some extent the starting coalition structure. In particular, we allow only for permutations of players such that the coalition $S \in \beta_S$ remains intact. Second, we will turn to a specific member of the class of PSSS, in particular the Equal Surplus Sharing Scheme ESSS characterized by $\lambda_i = \lambda_j,\ \forall i, j \in N$.

Let ρ be a permutation of the set N, *i.e.* a one-to-one function from N to N. For any partition function π and any coalition structure β_S, consider a permutation ρ under which coalition S is invariant (*i.e.* for each $j \in S,\ \rho(j) = i \in S$). In this case for each $\beta_S \in \mathcal{B}$ the corresponding coalition structure $\rho(\beta_S) \in \mathcal{B}$, in particular $\rho(S) = S$. Also, define the composition $\rho\pi$ as follows:

$$\rho\pi[\rho(\beta_S)] = \pi(\beta_S).$$

Axiom 3 : **Anonimity**

Consider a game in partition function form $\Gamma = (N, \mathcal{B}, \pi)$. A valuation function f^π satisfies anonimity if for any $i, j \in N$ such that $\rho(j) = i$ we have:

$$f_i^{\rho\pi}[\rho(\beta_S)] = f_j^{\pi}(\beta_S).$$

The anonymity property says that changing the order of the players, while leaving the coalition structure intact, does not change the valuation of the players. In fact, the outcome does depend on whether a player is an insider or outsider, but it does not depend on the exact position which a player takes among the group of coalition members or the group of singletons.

Proposition 5 : **Anonimity of ESSS**

The ESSS (for any non empty coalition S with cardinality s, $\lambda_i(\beta_S) = 1/s$) satisfies anonimity.

Proof:
1) $i \in S, \rho(j) = i \Rightarrow j \in S$. So $\rho(S) = S$.

$$v_i^{\rho\pi}[\rho(\beta_S)] = \rho\pi_i(\beta_{S\setminus\{i\}}) + \frac{1}{s}\left[\rho\pi_S(\beta_S) - \sum_{i\in\rho(S)} \rho\pi_i(\beta_{S\setminus\{i\}})\right] =$$

$$\pi_j(\beta_{S\setminus\{j\}}) + \frac{1}{s}\left[\pi_S(\beta_S) - \sum_{j\in S}\pi_j(\beta_{S\setminus\{j\}})\right] = v_j^{\pi}(\beta_S).$$

2) $i \notin S$. We consider $\rho(j) = i$, so $j \notin S$.

$$v_i^{\rho\pi}[\rho(\beta_S)] = \rho\pi_i(\rho(\beta_S)) = \pi_j(\beta_S) = v_j^{\pi}(\beta_S)$$

□

4.4 Axiomatization

Propositions 3-5 imply the following result.

Proposition 6 : **Axioms satisfied by ESSS**

Let $\Gamma(N, \mathcal{B}, \pi)$ *be a game in partition function form. For any coalition, the ESSS satisfies axioms coalitional efficiency, anonimity and additivity.*

Note that the assumption of equal weights is only necessary for anonimity. The other properties, coalitional efficiency and addivity, hold for the more general class of PSSS with arbitrary weights.

We now prove that if another sharing scheme or solution concept for a game in partition function form $\Gamma(N, \mathcal{B}, \pi)$, say f^{π}, satisfies axioms EFF, AN and ADD, then it coincides with the ESSVF, *i.e.* the PSSVF with equal weights.

Proposition 7 : **ESSS is the unique solution satisfying EFF, ADD and AN**

Let $\Gamma(N, \mathcal{B}, \pi)$ *be a game in partition function form. The ESSS for any coalition structure* $\beta_S \in \mathcal{B}$ *is the unique solution satisfying axioms coalitional efficiency, anonimity and additivity.*

Proof:
The result is proved in the same spirit of Shapley's characterization theorem, *i.e.* by using the linarity of the characteristic form games space. For this we introduce a technical definition, the so-called c-characteristic function. This allows us to specify the linear structure of the c-cooperative game space, *i.e.* the partition function form game space we are considering in this paper.

If $\mathcal{P}^*(N) = \{S \subseteq N, S \neq \emptyset\}$, given the partition function π, we will denote by π_c the *c-characteristic function* $\pi_c : \mathcal{P}^*(N) \times N \to \mathbb{R}$ defined for any $(S,i) \in \mathcal{P}^*(N) \times N$ as follows:

$$\pi_c(S,i) = \begin{cases} \pi_S(S,i) = \pi_S(\beta_S) & \text{if } i \in S, \\ \pi_i(S,i) = \pi_i(\beta_S) & \text{if } i \notin S. \end{cases}$$

Let us note that the c-characteristic function is constant on the set S. We call the pair (N, π_c) a c-cooperative game (which by definition coincides with the partition function form game $\Gamma = (N, \mathcal{B}, \pi)$) and denote by Π_c^n the space of all c-cooperative games with n players, *i.e.* the set of all possible π_c defined on $\mathcal{P}^*(N) \times N$. Let π_c^1 and π_c^2 be in Π_c^n and $\alpha \in \mathbb{R}$. Since for any $(S,i) \in \mathcal{P}^*(N) \times N$, $(\pi_c^1 + \pi_c^2)(S,i) = \pi_c^1(S,i) + \pi_c^2(S,i)$ and $(\alpha\pi_c^1)(S,i) = \alpha\pi_c^1(S,i)$, Π_c^n turns out to be a linear space.

The dimension of Π_c^n is $2^n - 1 + n2^{n-1} - n^2$.

We have one coalition structure of singletons $(\{1\}, ..., \{n\})$ and for every coalition S with $s \geq 2$ the number of possible coalition structures with a coalition of cardinality s is

$$\left(\begin{array}{c} n \\ s \end{array} \right)(n - s + 1)$$

and hence, the dimension is given by

$$n + \sum_{s=2}^{n} \left(\begin{array}{c} n \\ s \end{array} \right)(n - s + 1) = n + \sum_{s=2}^{n} \left(\begin{array}{c} n \\ s \end{array} \right) + n \sum_{s=2}^{n-1} \left(\begin{array}{c} n-1 \\ s \end{array} \right) =$$

$$2^n - 1 + n[2^{n-1} - n].$$

A basis is given by $\{\delta_T,\ T \subseteq N, T \neq \emptyset\} \cup \{\delta_{T,j},\ T \subset N, j \notin T\}$ where

$$\delta_T(S,i) = \begin{cases} 1 & \text{if } T = S \text{ and } i \in T, \\ 0 & \text{otherwise.} \end{cases}$$

$$\delta_{T,j}(S,i) = \begin{cases} 1 & \text{if } T = S \text{ and } i = j \notin T, \\ 0 & \text{otherwise.} \end{cases}.$$

Let $\pi \in \Pi_c^n$ be a c-characteristic function. Then for any $(S,i) \in \mathcal{P}^*(N) \times N$ we have

$$\pi_c(S,i) = \sum_{T \subset N} \{c_T \delta_T(S,i) + \sum_{j \notin T} c_{T,j} \delta_{T,j}(S,i)\} + c_N \delta_N(S,i)$$

with $c_T, c_{T,j} \in \mathbb{R}$.

Let $\Gamma = (N, \mathcal{B}, \pi)$ and π be a partition function; for any coalition structure $\beta_S \in \mathcal{B}$ define π_c the c-characteristic function associated with the partition function π for the game Γ. In the following, we will omit the subscript c, the argument will specify the function we are considering. Let f^π be a solution concept for the partition function π, *i.e.*

$$f^\pi : \beta_S = (S, \{j_1\}, ..., \{j_{n-s}\}) \in \mathcal{B} \mapsto (f_1^\pi(\beta_S), ..., f_n^\pi(\beta_S)) \in \mathbb{R}^n.$$

For any $\beta_S \in \mathcal{B}$, we can also consider the solution f^π defined on the set $\mathcal{P}^*(N) \times N$ as follows

$$f^\pi(S, i) = \begin{cases} f_i^\pi(\beta_S) & \text{if } i \in S, \\ f_j^\pi(\beta_S) = \pi_j(\beta_S) & \text{if } j \notin S. \end{cases}$$

Let us suppose that the solution concept f^π satisfies axioms coalitional efficiency, anonimity and additivity. By using additivity we have:

$$f^\pi = f^{c_N \delta_N} + \sum_{T \subset N} f^{c_T \delta_T + \sum_{j \notin T} c_{T,j} \delta_{T,j}}$$

and since also the ESSS ν^π satisfies additivity,

$$\nu^\pi = \nu^{c_N \delta_N} + \sum_{T \subset N} \nu^{c_T \delta_T + \sum_{j \notin T} c_{T,j} \delta_{T,j}}.$$

It is sufficient to prove that for any $(S, i) \in \mathcal{P}^*(N) \times N$, we have

$$f^{c_N \delta_N}(S, i) = \nu^{c_N \delta_N}(S, i) \tag{1}$$

and for any coalition $T \subset N$, for any $(S, i) \in \mathcal{P}^*(N) \times N$ and $j \notin T$, we have

$$f^{c_T \delta_T + \sum_{j \notin T} c_{T,j} \delta_{T,j}}(S, i) = \nu^{c_T \delta_T + \sum_{j \notin T} c_{T,j} \delta_{T,j}}(S, i). \tag{2}$$

Since for any $(S, i) \in \mathcal{P}^*(N) \times N$ the function $\delta_N(S - \{i\}, i) = 0$ by definition of ESSS, we have

$$\nu^{c_N \delta_N}(S, i) = c_N \delta_N(S - \{i\}, i) + 1/s[c_N \delta_N(S, i) - \sum_{i \in S} c_N \delta_N(S - \{i\}, i)] = c_N 1/n$$

if $S = N$ and zero otherwise.

By EFF we have $\sum_{i \in S} f^{c_N \delta_N}(S, i) = c_N \delta_N(S, i) = c_N$ if $S = N$ and zero otherwise. Let us consider a permutation $\rho : N \to N$ such that $\rho(S) = S$ and $\rho(j) = i$; define $\rho c_N \delta_N(\rho(S), i) = c_N \delta_N(S, j)$. By using AN we have $f^{\rho c_N \delta_N}(\rho(N), i) = f^{c_N \delta_N}(N, j)$ so that for any $i, j \in N$ the values coincide $f^{c_N \delta_N}(N, i) = f^{c_N \delta_N}(N, j)$. This implies that $\sum_{i \in N} f^{c_N \delta_N}(N, i) = n f^{c_N \delta_N}(N, i) = c_N$. Equality (1) is therefore proven.

Fix now a coalition $T \subset N$. We have to prove equality (2) by distinguishing four cases.

$2_a)$ $S = T, i = j \notin S = T$:

By definition of ESSS being $i \notin S$ we have $\nu^{c_S\delta_S+\sum_{i\notin S} c_{S,i}\delta_{S,i}}(S,i) = c_{S,i}$. But then, by using EFF with $i \notin S$, we have

$$f^{c_T\delta_T}(S,i) + \sum_{j\notin T} f^{c_{T,j}\delta_{T,j}}(S,i) = c_T\delta_T(S,i) + \sum_{j\notin T} c_{T,j}\delta_{T,j}(S,i) =$$

$$0 + \sum_{j\notin S} c_{S,j}\delta_{S,j}(S,i) = c_{S,i}.$$

$2_b)$ $S = T, i \in S, j \notin S = T$:

Since $i \in S$, $\delta_{T,j}(S,i) = 0$ and since $j \notin T$ by definition of ESSS, we have that $\nu^{c_T\delta_T+\sum_{j\notin T} c_{T,j}\delta_{T,j}}(S,i) = \nu^{c_T\delta_T}(S,i) = c_S 1/s$. Analogously, $f^{c_T\delta_T+\sum_{j\notin T} c_{T,j}\delta_{T,j}}(S,i) = f^{c_S\delta_S}(S,i)$, and by EFF we have $\sum_{i\in S} f^{c_S\delta_S}(S,i) = c_S\delta_S(S,i) = c_S$. By using AN we have $f^{c_S\delta_S}(S,i) = f^{c_S\delta_S}(S,j)$. This implies that $\sum_{i\in S} f^{c_S\delta_S}(S,i) = sf^{c_S\delta_S}(S,i) = c_S$.

$2_c)$ $T = S-\{i\}, i = j \in S, j \notin T$:

We prove that

$$f^{c_{S-\{i\}}\delta_{S-\{i\}}+\sum_{i\notin S-\{i\}} c_{S-\{i\},i}\delta_{S-\{i\},i}}(C,h) = \nu^{c_{S-\{i\}}\delta_{S-\{i\}}+\sum_{i\notin S-\{i\}} c_{S-\{i\},i}\delta_{S-\{i\},i}}(C,h)$$

for any $C \subseteq N$ and $h \in N$.

If $h \notin C$, by EFF equality (2) is satisfied because both members are equal to

$$\gamma(C,h) = [c_{S-\{i\}}\delta_{S-\{i\}} + \sum_{i\notin S-\{i\}} c_{S-\{i\},i}\delta_{S-\{i\},i}](C,h).$$

For $h \in C$, the function $\gamma(C,h)$ is zero if $C \neq S-\{i\}$ and equality (2) holds because both members are 0. Let us suppose $C = S-\{i\}$ and $h \in S-\{i\}$. So

$$\nu^\gamma(S-\{i\},h) = \gamma_h(S-\{i\}-\{h\},h) + 1/(s-1)[\gamma_{S-\{i\}}(S-\{i\},h) -$$

$$\sum_{h\in S-\{i\}} \gamma_h(S-\{i\}-\{h\},h)] = c_{S-\{i\}}1/(s-1).$$

But then, $f^\gamma(S-\{i\},h) = f^{c_{S-\{i\}}\delta_{S-\{i\}}}(S-\{i\},h)$. By using EFF

$$\sum_{h\in S-\{i\}} f^{c_{S-\{i\}}\delta_{S-\{i\}}}(S-\{i\},h) = c_{S-\{i\}}\delta_{S-\{i\}}(S-\{i\},h) = c_{S-\{i\}}$$

and by using AN

$$f^{c_{S-\{i\}}\delta_{S-\{i\}}}(S-\{i\},h) = c_{S-\{i\}}1/(s-1).$$

$2_d)$ $T \neq S, T \neq S-\{i\}, i \in S \cap T, j \notin T$:

In this case the function $\gamma(S,i) = [c_T\delta_T + \sum_{j\notin T} c_{T,j}\delta_{T,j}](S,i) = 0$, hence equality (2) holds because both members are 0. □

5 Conclusion

In this paper, we have introduced a new sharing scheme, the Proportional Surplus Sharing Scheme PSSS for the gains from cooperation in games with externalities that suffer from free-rider problems and where the formation of the grand coalition cannot be taken for granted. We focused on partition function games in which coalition structures consist of one genuine coalition plus singletons and for which the underlying economic problem is associated with positive externalities of coalition formation, *i.e.* games in which outsiders to a coalition benefit from the expansion of the coalition. Prominent examples in this class of games are the provision of public goods like international cooperation between countries on issues of security, environment, customs or monetary unions, and cooperation between private companies on matters like R&D, procurement, output or price coordination.

We showed that the PSSS achieves the maximum aggregate welfare subject to the constraint that coalitions have to be stable in the sense of d'Aspremont et al. (1983). More precisely, PSSS attains the highest possible aggregate welfare among the set of coalitions that can potentially be stabilized internally. Potential internal stability PIS means that coalitions generate sufficient surplus to compensate for the free riding claims of their members. The grand coalition may or may not be a member of this set of PIS coalitions. The sharing scheme is flexible because, irrespective of the weights, the maximum welfare result holds.

The other important contribution of this paper was that we provided a complete characterization of a particular member of the class of PSSS. It turns out that, much in the spirit of charaterizations of the standard Shapley value in cooperative game theory, coalitional efficiency, additivity and anonimity, completely characterize the Equal Surplus Sharing Scheme ESSS, which is a PSSS with equal sharing weights.

Hence, we improved upon the existing literature on coalition formation in the context of externalities by departing from the assumption of symmetric agents, establishing existence of non-trivial stable coalitions under mild conditions, characterize the solution concept in terms of well established axioms and showed optimality subject to stability. For future research, at least two possible extensions come to our mind. First, our approach could be generalized to allow for the co-existence of several non-trivial coalitions. Of course, multiple coalitions would complicate the analysis tremendously because outside options or threat point payoffs are not straightforwardly defined any longer as threat points are mutually depended and linked. Second, for many of our results, the PSSS leaves the choice of surplus sharing weights open. Endogenizing the value of these weights, which may be interpreted as bargaining power, in games with heterogeneous players and externalities seems an interesting topic for further research.

References

Aumann, R.J. and J.H. Drèze (1974), Cooperative Games with Coalition Structures, *International Journal of Game Theory* **3**, 217–237.

Barrett, S. (1997), Heterogeneous International Agreements, in: Carraro, C. (ed.), *International Environmental Negotiations: Strategic Policy Issues*, (Edward Elgar, Cheltenham, UK), pp. 9–25.

Bloch, F. (2003), Non-Cooperative Models of Coalition Formation in Games with Spillovers, in: Carraro, C. (ed.), *Endogenous Formation of Economic Coalitions*, (Edward Elgar, Cheltenham, UK), chapter 2, pp. 35–79.

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

Botteon, M., and Carraro, C. (1997), Burden-sharing and coalition stability in environmental negotiations with asymmetric countries, in: C. Carraro (Ed.), *International environmental negotiations: Strategic policy issues* (Edward Elgar, Cheltenham, UK), chapter 3, pp. 26-55).

d'Aspremont, C., A. Jacquemin, J. J. Gabszewicz and J. A. Weymark (1983), Pn the Stability of Collusive Price Leadership, *Canadian Journal of Economics* **16**(1), 17–25.

de Clippel, G. and R. Serrano (2008), Marginal Contributions and Externalities in the Value. *Econometrica* **76**(6), 1413–1436.

Deneckere, R. and C. Davidson (1985), Incentives to Form Coalitions with Bertrand Competition, *Rand Journal of Economics* **16**, 473–486

Faigle, U. and Grabisch, M. (2011), Values for Markovian Coalition Processes, *Economic Theory*, forthcoming.

Fuentes-Albero, C., and Rubio, S.J. (2010), Can international environmental cooperation be bought? *European Journal of Operational Research* **202**, 255–264.

Grabisch, M. and Funaki, Y. (2008), A Coalition Formation Value for Games with Externalities, mimeo, Documents de travail du Centre d'Economie de la Sorbonne, Paris, http://halshs.archives-ouvertes.fr/docs/00/34/47/97/PDF/B08076.pdf

Hafalir, I.E. (2007), Efficiency in Coalition Games with Externalities, *Games and Economic Behavior* **61**, 242–258.

Kohler, M. (2002), Coalition Formation in International Monetary Policy Games, *Journal of International Economics* **56**, 371–385.

Kolpin, V. (1996), Multi-Product Serial Cost Sharing: An Incompatibility with the Additivity Axiom, *Journal of Economic Theory* **69**, 227–233.

Macho-Stadler, I., Pérez-Castrillo, D. and Wettstein, D. (2007) Sharing the Surplus: an Extension of the Shapley Value for Environments with Externalities, *Journal of Economic Theory* **135**, 339–356.

Maskin, E. (2003), Bargaining, Coalitions, and Externalities. Mimeo, paper presented at the EEA-ESEM Meeting in Stockholm, August 20-24, 2003.

McGinty, M. (2007), International environmental agreements among asymmetric nations, *Oxford Economic Papers* **59**, 45–62.

Meyerson, R.B. (1977), Values of Games in Partition Function Form, *International Journal of Game Theory* **6**, 23–31.

Poyago-Theotokay, J. (1995), Equilibrium and Optimal Size of A Research Joint Venture in an Oligopoly with Spillovers, *The Journal of Industrial Economics* **43**, 209–226.

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

Ray, D. and Vohra, R. (2001), A Theory of Endogenous Coalition Structures, *Journal of Political Economy* **109**(6), 1355–1384.

Sandler T. (2004), *Global Collective Action* (Cambridge University Press, Cambridge UK).

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

Weikard, H.-P., M. Finus and J. C. Altamirano-Cabrera (2006), The Impact of Surplus Sharing on the Stability of International Climate Coalitions, *Oxford Economic Papers* **58**, 209–232.

Weikard, H.-P. (2007), Cartel Stability under an Optimal Sharing Rule, *Manchester School* **77**(5), 575–593.

Yi, S.-S. (1996), Endogenous Formation of Customs Unions under Imprefect Competition: Open Regionalism is Good. *Journal of International Economics* **41**, 153–177.

Yi, S.-S. (1997), Stable Coalition Structures with Externalities, *Games and Economic Behavior* **20**(2), 201–237.

Yi, S.-S. (2003), Endogenous Formation of Economic Coalitions: A Survey of the Partition Function Approach, in: Carraro, C. (ed.), *Endogenous Formation of Economic Coalitions*, (Edward Elgar, Cheltenham, UK), chapter 3, pp. 80–127.

Yi, S.-S. and H. Shin (2000), Endogenous Formation of Research Coalitions with Spillovers, *International Journal of Industrial Organization* **18**, 229–256.

Appendix

Table A.1 displays a four player example with a partition function that exhibits positive externalities. The table displays the coalition structure (column 1), the worth to coalition S (column 2) and to non-members of S (column 3-6) and the aggregate worth to all players (column 7), the sum of free-rider payoffs (column 8), the surplus to the coalition (column 9; Definition 4) which is the difference between column 8 and 2 and indicates coalitions that are (not) potentially internally stable coalitions with 1 (0) (column 10).

Table A.1: Partition Function

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
β_S	π_S	π_a	π_b	π_c	π_d	π_N	$\pi_S(\beta_{S\setminus\{j\}})$	$\sigma(\beta_S)$	Σ^{PIS}
$(\{a\},\{b\},\{c\},\{d\})$	-	0	0	0	0	0	-	-	1
$(\{a,b\},\{c\},\{d\})$	1	-	-	1	1	3	0	1	1
$(\{a,c\},\{b\},\{d\})$	1	-	$2\frac{1}{3}$	-	$3\frac{1}{3}$	$6\frac{2}{3}$	0	1	1
$(\{a,d\},\{b\},\{c\})$	$2\frac{1}{3}$	-	1	2	-	$5\frac{1}{3}$	0	$2\frac{1}{3}$	1
$(\{b,c\},\{a\},\{d\})$	3	$3\frac{1}{3}$	-	-	1	$7\frac{1}{3}$	0	3	1
$(\{b,d\},\{a\},\{c\})$	1	$3\frac{1}{3}$	-	2	-	$6\frac{1}{3}$	0	1	1
$(\{c,d\},\{a\},\{b\})$	1	1	6	-	-	8	0	1	1
$(\{a,b,c\},\{d\})$	9	-	-	-	$4\frac{1}{3}$	$13\frac{1}{3}$	$6\frac{2}{3}$	$2\frac{1}{3}$	1
$(\{a,b,d\},\{c\})$	$5\frac{1}{3}$	-	-	3	-	$8\frac{1}{3}$	$5\frac{1}{3}$	0	1
$(\{a,c,d\},\{b\})$	$5\frac{1}{3}$	-	7	-	-	$12\frac{1}{3}$	$6\frac{1}{3}$	-1	0
$(\{b,c,d\},\{a\})$	8	$4\frac{1}{3}$	-	-	-	$12\frac{1}{3}$	9	-1	0
$(\{a,b,c,d\})$	$17\frac{2}{3}$	-	-	-	-	$17\frac{2}{3}$	$18\frac{2}{3}$	-1	0

We display valuations for the Shapley value in Table A.2. We use the extended version of the Shapley value as defined in Aumann and Drèze (1974), meaning that these values are computed not only for the grand coalition but for every non-empty coalition. The numbers for the Shapley value are computed according to the following

formula:

$$\forall \beta_S \in \mathcal{B},\ \forall i \in S: \qquad v_i^{Shp}(\beta_S) = \sum_{T \subseteq S} \frac{(t-1)!(s-t)!}{t!} \left[\pi_T(\beta_T) - \pi_{T\setminus\{i\}}(\beta_{T\setminus\{i\}})\right].$$

Table A.2: Valuation Function for the Shapley Value

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
β_S	v_a^{Shp}	v_b^{Shp}	v_c^{Shp}	v_d^{Shp}	v_N^{Shp}	Σ^{IS}	Σ^{ES}	Σ^{S}
$(\{a\},\{b\},\{c\},\{d\})$	0	0	0	0	0	1	0	0
$(\{a,b\},\{c\},\{d\})$	$\frac{1}{2}$	$\frac{1}{2}$	1	1	3	1	0	0
$(\{a,c\},\{b\},\{d\})$	$\frac{1}{2}$	$2\frac{1}{3}$	$\frac{1}{2}$	$3\frac{1}{3}$	$6\frac{2}{3}$	1	0	0
$(\{a,d\},\{b\},\{c\})$	$1\frac{1}{6}$	1	2	$1\frac{1}{6}$	$5\frac{1}{3}$	1	0	0
$(\{b,c\},\{a\},\{d\})$	$3\frac{1}{3}$	$1\frac{1}{2}$	$1\frac{1}{2}$	1	$7\frac{1}{3}$	1	0	0
$(\{b,d\},\{a\},\{c\})$	$3\frac{1}{3}$	$\frac{1}{2}$	2	$\frac{1}{2}$	$6\frac{1}{3}$	1	0	0
$(\{c,d\},\{a\},\{b\})$	1	6	$\frac{1}{2}$	$\frac{1}{2}$	8	1	0	0
$(\{a,b,c\},\{d\})$	$2\frac{1}{3}$	$3\frac{1}{3}$	$3\frac{1}{3}$	$4\frac{1}{3}$	$13\frac{1}{3}$	0	1	0
$(\{a,b,d\},\{c\})$	2	$1\frac{1}{3}$	3	2	$8\frac{1}{3}$	0	0	0
$(\{a,c,d\},\{b\})$	2	7	$1\frac{1}{3}$	2	$12\frac{1}{3}$	0	1	0
$(\{b,c,d\},\{a\})$	$4\frac{1}{3}$	3	3	2	$12\frac{1}{3}$	0	1	0
$(\{a,b,c,d\})$	4	5	5	$3\frac{2}{3}$	$17\frac{2}{3}$	0	1	0

Valuations are computed applying the formula above to the partition function in Table A.1.

Table A.3: Valuation Function for the Proportional Surplus Sharing Scheme

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
β_S	v_a^{PSSS}	v_b^{PSSS}	v_c^{PSSS}	v_d^{PSSS}	v_N^{PSSS}	Σ^{IS}	Σ^{ES}	Σ^{S}
$(\{a\},\{b\},\{c\},\{d\})$	0	0	0	0	0	1	0	0
$(\{a,b\},\{c\},\{d\})$	$\frac{1}{2}$	$\frac{1}{2}$	1	1	3	1	0	0
$(\{a,c\},\{b\},\{d\})$	$\frac{1}{2}$	$2\frac{1}{3}$	$\frac{1}{2}$	$3\frac{1}{3}$	$6\frac{2}{3}$	1	0	0
$(\{a,d\},\{b\},\{c\})$	$1\frac{1}{6}$	1	2	$1\frac{1}{6}$	$5\frac{1}{3}$	1	1	1
$(\{b,c\},\{a\},\{d\})$	$3\frac{1}{3}$	$1\frac{1}{2}$	$1\frac{1}{2}$	1	$7\frac{1}{3}$	1	0	0
$(\{b,d\},\{a\},\{c\})$	$3\frac{1}{3}$	$\frac{1}{2}$	2	$\frac{1}{2}$	$6\frac{1}{3}$	1	1	1
$(\{c,d\},\{a\},\{b\})$	1	6	$\frac{1}{2}$	$\frac{1}{2}$	8	1	1	1
$(\{a,b,c\},\{d\})$	$4\frac{1}{9}$	$3\frac{1}{9}$	$1\frac{7}{9}$	$4\frac{1}{3}$	$13\frac{1}{3}$	1	1	1
$(\{a,b,d\},\{c\})$	$3\frac{1}{3}$	1	3	1	$8\frac{1}{3}$	1	1	1
$(\{a,c,d\},\{b\})$	$\frac{2}{3}$	7	$1\frac{2}{3}$	3	$12\frac{1}{3}$	0	1	0
$(\{b,c,d\},\{a\})$	$4\frac{1}{3}$	$5\frac{2}{3}$	$1\frac{2}{3}$	$\frac{2}{3}$	$12\frac{1}{3}$	0	1	0
$(\{a,b,c,d\})$	$4\frac{1}{12}$	$6\frac{3}{4}$	$2\frac{3}{4}$	$4\frac{1}{12}$	$17\frac{2}{3}$	0	1	0

Valuations are computed applying Definition 4 to the partition function in Table A.1 and assuming equal weights, $\lambda_i(\beta_S) = \frac{1}{s}\ \forall i \in S$ and $\forall \beta_S$.