Expectation Formation Rules and the Core of Partition Function Games

Francis Bloch* Anne van den Nouweland†

May 13, 2013

Abstract

This paper proposes axiomatic foundations of expectation formation rules, by which deviating players anticipate the reaction of external players in a partition function game. The axioms single out the projection rule among the rules that depend on the current partition and the pessimistic rule among the ones that are independent of the current partition. This analysis suggests that the projection core and the pessimistic core are natural candidates to study the stability of games in partition function form, and we compute these cores in two standard applications of coalition formation with externalities, namely cartels and public goods.

`JEL Classification Numbers: C71, C70`

`Keywords: Partition Function Games, Core, Expectation Formation, Axiomatization`

*Corresponding author. Department of Economics, Ecole Polytechnique, 91128 Palaiseau, France. Email francis.bloch@polytechnique.edu.

†Department of Economics 435 PLC, 1285 University of Oregon, Eugene OR 97403-1285, USA. Email annev@uoregon.edu.

1 Introduction

The objective of this paper is to provide axiomatic foundations for extensions of the core to games in partition function form. It is well known that, if one moves beyond the highly competitive, zero-sum game environment of van Neumann and Morgenstern (1944), the worth of a coalition cannot be defined independently of the coalition structure formed by other players. The natural description of a cooperative environment is then a game in partition function form (Thrall and Lucas (1963)) specifying for each coalition structure and each coalition embedded in that coalition structure, the worth that the coalition can achieve. Ray (2007) contains a thorough discussion of the difference between partition function games and coalitional games, and references to the early literature on partition functions.

Unfortunately, in games in partition function form, the dominance relation which supports the core cannot be defined unambiguously. When a coalition of players deviates, the payoff they expect to obtain depends on the way external players react to the deviation. This ambiguity has long been recognized – at least since Aumann (1967) –, and various definitions of the core have been proposed corresponding to different specifications of the expectations of deviating players on the reaction of external players. For example, Hart and Kurz (1983) describe the α and β cores, based on pessimistic beliefs where players expect external players to organize in such a way that they minimize the payoffs of deviating players, and the γ and δ cores, where players anticipate that coalitions which have been left by some members of the deviating group either disintegrate into singletons, or stick together.[1] Chander and Tulkens (1997) and de Clippel and Serrano (2008) focus attention on a model where deviating coalitions expect all other players to remain singletons, whereas Shenoy (1979) assumes that deviating players are optimistic and anticipate that external players organize in order to maximize their payoffs.

Definitions of the core of partition function games proposed in the literature are thus based on ad hoc assumptions on the reaction of external players to the deviation. By contrast, our objective in this paper is to ground the expectations of deviating players on axioms, and derive the core of a partition function game on the basis of properties satisfied by the expectation

[1] The γ model finds its roots in Von Neumann and Morgenstern (1944) who discuss a game of coalition formation among three agents which requires unanimity and is equivalent to the γ game.

formation rule. We first propose a set of axioms that pertains to the relation between the current partition and the expectations formed by deviating players. An expectation formation rule is independent of the original partition or independent of the position of deviating players in the original partition if players do not tie their expectations to the current state. It is instead responsive if different partitions of external players always give rise to different expectations. The second set of axioms deals with the consistency of expectations among groups of players. Path independence states that, when a coalition $S \cup T$ deviates, the expectations they form on the reaction of external players is the same, whether S deviates first and T second, or T deviates first and S second. Subset consistency specifies that the expectations formed by a group of players S and by any subset T of S must give rise to the same organization of players in the complement of the larger set S. Coherence of expectations introduces a consistency condition between the expectations formed by S and its complement. Finally, we define preservation of superadditivity as the property which guarantees that, if the underlying partition function is superadditive, then for any partition, the coalitional function which results from the expectation formation rule is also superadditive.

We analyze which of the commonly used expectation formation rules satisfy these axioms, and characterize the *projection rule* (by which players anticipate that external players form a coalition structure which is the projection of the current coalition structure) as the only expectation formation rule that satisfies responsiveness and subset consistency or subset consistency, independence of position of deviating players in the original partition, and coherence of expectations. Hence, we believe that the projection rule (which is equivalent to the δ rule proposed by Hart and Kurz (1983)) has a solid axiomatic basis and is a natural candidate for an expectation formation rule. If instead of responsiveness to the current partition, we require independence of the current partition, the only rules that satisfy subset consistency are *exogenous rules* where deviating players anticipate external players to organize according to the projection of an exogenous partition $\mathcal{M}$. Notice in particular that if $\mathcal{M}$ is a partition of singletons, the $\mathcal{M}$-exogenous rule corresponds to the γ rule of Chander and Tulkens (1997) or the externality-free rule of de Clippel and Serrano (2008) whereas if $\mathcal{M}$ is the partition formed by the grand coalition, the $\mathcal{M}$-exogenous rule specifies that agents anticipate external players to form a single component. Finally we note that the *pessimistic rule* (the α rule) is the only expectation formation rule which satisfies preservation of superadditivity.

Equipped with a description of the reaction of external players to a deviation, we define the core of a partition function game by deriving the coalitional game generated by the expectation formation rule. To define the *projection core*, we first need to compute the coalition structure $\mathcal{N}^*$ that maximizes the sum of payoffs of all agents, and then construct the coalitional game where deviating players in S anticipate external players to form the projection of $\mathcal{N}^*$. The projection core always lies between the pessimistic core (the largest core) and the optimistic core (the smallest core). In fact, we remark that if the game is superadditive and has positive externalities, the projection core coincides with the optimistic core whereas if the game is superadditive and has negative externalities, the projection core coincides with the pessimistic core.

We compute the projection core and the optimistic and pessimistic cores in two standard applications of games of coalition formation with externalities. In a game of cartel formation with a fixed cost, similar to the game considered by Bloch (1996) and Ray and Vohra (1999), we show that the projection core is identical to the optimistic core, is nonempty if and only if the fixed cost is higher than a lower bound, and that the pessimistic core is always nonempty. In a public good game with spillovers similar to the game proposed by Ray and Vohra (2001), we note that the projection core is identical to the optimistic core, is nonempty if and only if spillovers are lower than an upper bound, and that the pessimistic core is always nonempty.

To the best of our knowledge, our paper represents the first attempt to axiomatize the reaction of external players to a deviation in order to define the core of partition function games. However, the need to specify the partition of external players also appears in studies of extensions of the Shapley value to partition function games. Starting with Myerson (1977), several extensions of the Shapley value to partition function games have been proposed. Recently, Macho-Stadler, Perez-Castrillo and Wettstein (2007) have proposed an axiomatization based on the classical axioms of Shapley. De Clippel and Serrano (2008) base their value on axioms of marginality and monotonicity. Dutta, Ehlers and Kar (2010) extend the axioms of consistency and the potential approach to partition function games. While the axioms we discuss in the current paper are applied to a different object than the axioms studied in the context of the Shapley value, there are clear similarities between our approaches. In order to use the potential approach, Dutta, Ehlers and Kar (2010) need to define restrictions of partition function games after one player leaves. They propose axioms on restriction operators, including a path independence axiom which guarantees that the restricted games do not depend on the order in which players leave. Implicitly, their axioms embody conditions on the partition formed after a player leaves. By contrast, our axioms apply directly to expectation formation rules. Hence, their axiomatizations and ours are complementary.

The rest of the paper is organized as follows. We present our model of partition function games and expectation formation rules in the next section. Section 3 is devoted to the description of axioms on expectation formation rules. Section 4 contains the core of our analysis with the axiomatizations of the projection and exogenous rules and a discussion of preservation of superadditivity. We discuss the construction of the core of partition function games generated by expectation formation rules in Section 5 and present two applications to standard games of coalition formation with externalities in Section 6. Section 7 concludes and proposes directions for future research.

2 The Model

2.1 Partition function games

We consider a set N of players with cardinality $n \geq 3$. A partition on N is a collection of pairwise disjoint, nonempty subsets of N covering N. Let $\Pi(N)$ be the set of all partitions on N, with typical element $\mathcal{N}$. Similarly, for any subset S of N, we denote by $\Pi(S)$ the set of all partitions on S with typical element $\mathcal{S}$. The partition of S formed only of singletons is denoted $\underline{\underline{S}} = \{\{i\} \mid i \in S\}$ and the partition of S formed only by the set S is denoted $\overline{\overline{S}} = \{S\}$. For any set S, S^c denotes the complement of S in N. Given a set S and a partition $\mathcal{S}$ of S and a subset T of S, we let $\mathcal{S}|_T$ denote the projection of $\mathcal{S}$ onto T, i.e. the partition $\mathcal{T}$ of T such that i and j belong to the same block in $\mathcal{T}$ if and only if they belong to the same block in $\mathcal{S}$.

We suppose that the strategic situation faced by the agents is captured by a *TU game in partition function form.* Partition function games, introduced by Thrall and Lucas (1963), generalize coalitional games by allowing for externalities across coalitions. They arise naturally in environments where players can form binding agreements, and cooperate inside coalitions but compete across coalitions (see Ray (2007)). Formally, a partition function v associates to each partition $\mathcal{N}$ and each block $S \in \mathcal{N}$ a positive number $v(S, \mathcal{N})$ specifying the *worth of coalition S in partition $\mathcal{N}$*. Notice that a partition function only assigns worths to those subsets which are blocks in $\mathcal{N}$. If S does not belong to $\mathcal{N}$, then $v(S, \mathcal{N})$ is not defined. A partition function v is *superadditive* if for all $\mathcal{N} \in \Pi$, for all $S, T \in \mathcal{N}$,

$$v(S \cup T, \mathcal{N} \setminus \{S, T\} \cup \{S \cup T)\}) \geq v(S, \mathcal{N}) + v(T, \mathcal{N}).$$

It v has *positive externalities* if for all $\mathcal{N} \in \Pi$, for all $S \in \mathcal{N}$, and all $T, U \in \mathcal{N},\ T, U \subseteq S^c$,

$$v(S, \mathcal{N} \setminus \{T, U\} \cup \{T \cup U)\}) \geq v(S, \mathcal{N})$$

and *negative externalities* if for all $\mathcal{N} \in \Pi$, for all $S \in \mathcal{N}$, and all $T, U \in \mathcal{N},\ T, U \subseteq S^c$,

$$v(S, \mathcal{N} \setminus \{T, U\} \cup \{T \cup U)\}) \leq v(S, \mathcal{N}).$$

2.2 Coalitional games

A *TU game in coalitional function form* associates a real number to any nonempty subset of N. Formally, for any $S \subseteq N$, $S \neq \emptyset$, $w(S) \in \Re_+$ denotes the *worth of coalition S.* A coalitional game is superadditive if the worth of the union of two disjoint coalitions is greater than the sum of the worths. This property is justified by the fact that members of the two merging coalitions can always reproduce the behavior they adopted when the coalitions where separate, and can in addition benefit from cooperating after merging the two coalitions. A coalitional function w is *superadditive* if for all S, T such that $S \cap T = \emptyset$,

$$w(S \cup T) \geq w(S) + w(T).$$

2.3 Expectation formation rules

In a partition function game, when a coalition of players S contemplates deviating from a partition $\mathcal{N}$, they need to form expectations on the reaction of external players to their deviation. We define *expectation formation rules* assigning to every deviating coalition an expectation over the coalition formed by external players. We assume that coalitions have *deterministic expectations*, and that expectations may depend on the current partition $\mathcal{N}$, on the partition formed by deviating players $\mathcal{S}$ and on the partition function v.

Definition 2.1 *An* expectation formation rule *is a mapping f associating a partition $f(S, \mathcal{S}, \mathcal{N}, v)$ of S^c with each coalition S, partition $\mathcal{S}$ of S, partition $\mathcal{N}$ of N, and partition function v.*[2]

[2] The expectation formation rule is not defined for $S = N$ and $S = \emptyset$.

2.4 Generating coalitional functions from partition functions

For any expectation formation rule f and partition $\mathcal{N} \in \Pi(N)$, we generate a coalitional function $w_f^{\mathcal{N}}$ from the partition function v by assuming that external players react to a deviation according to f. When coalition S deviates, reorganizes itself into a partition $\mathcal{S}$, and expects external players to react according to f, it obtains an expected worth of

$$v(S, \mathcal{S}, f(S, \mathcal{S}, \mathcal{N}, v)).$$

Clearly, this expression depends on $\mathcal{S}$, the partition formed by members. In order to obtain a coalitional function, we assume that members of S reorganize into the partition $\mathcal{S}$ which maximizes the sum of worths of the coalition and define

$$w_f^{\mathcal{N}}(S) = \max_{\mathcal{S} \in \Pi(S)} \sum_{T \in \mathcal{S}} v(T, \mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N}, v)). \tag{1}$$

Notice that, in general, the coalitional function $w_f^{\mathcal{N}}$ is indexed by the current partition $\mathcal{N}$. However, $w_f^{\mathcal{N}}$ is *not* a partition function, as it assigns worths to *all* subsets S of N, including subsets which are not blocks in $\mathcal{N}$.

2.5 Some rules of expectation formation

Starting with von Neumann and Morgenstern (1944), game theorists have made different assumptions about reactions of external players in order to reduce partition function games to coalitional function games and apply classical solution concepts to predict the formation of coalitions and distribution of coalitional payoffs.

The disintegration rule Von Neumann and Morgenstern (1944) suppose that coalitions can only be formed by unanimous agreement of their members, resulting in an expectation formation rule where deviating players expect the coalitions they leave to disintegrate into singletons. In this model (also labeled the γ rule by Hart and Kurz (1983)), for any $T \in \mathcal{N}$, such that $T \cap S \neq \emptyset$, $T \setminus S$ disintegrates into $\overline{T \setminus S}$ in $f(S, \mathcal{S}, \mathcal{N}, v)$ and for any $T \in \mathcal{N}$ such that $T \cap S = \emptyset$, T remains in $f(S, \mathcal{S}, \mathcal{N}, v)$.

The projection rule Hart and Kurz (1983) introduce the δ model of coalition formation, where coalitions are formed by all players announcing the same coalition. This results in an expectation rule where players expect the coalitions that they leave to remain together. Hence, the expectation rule is given by $f(S, \mathcal{S}, \mathcal{N}, v) = \mathcal{N}|_{S^c}$.

$\mathcal{M}$-*Exogenous rules* An exogenous rule is indexed by a partition $\mathcal{M}$ of N. Players in S expect that external players organize according to the projection of $\mathcal{M}$ onto S^c: $f(S, \mathcal{S}, \mathcal{N}, v) = \mathcal{M}|_{S^c}$. Two special exogenous rules are the $\underline{N}$-exogenous rule, where players anticipate that all external players will form singletons (Tulkens and Chander (1997) and de Clippel and Serrano (2008)), and the $\overline{N}$-exogenous rule, where players anticipate that all external players join in a single coalition S^c.

The optimistic rule According to the optimistic rule, proposed by Shenoy (1979), players expect external members to select the[3] partition which maximizes the payoff of the players in S:

$$f(S, \mathcal{S}, \mathcal{N}, v) = \text{argmax}_{\mathcal{S}^c \in \Pi(S^c)} \sum_{T \in \mathcal{S}} v(T, \mathcal{S} \cup \mathcal{S}^c).$$

The pessimistic rule In the pessimistic rule, inspired by Aumann's (1967)'s definition of the α-core, and discussed by Hart and Kurz (1983), players expect external players to select the partition which minimizes the payoff of the players in S:

$$f(S, \mathcal{S}, \mathcal{N}, v) = \text{argmin}_{\mathcal{S}^c \in \Pi(S^c)} \sum_{T \in \mathcal{S}} v(T, \mathcal{S} \cup \mathcal{S}^c).$$

The max rule In the max rule, players expect external players to select a partition which maximizes the payoff of the players in S^c:

$$f(S, \mathcal{S}, \mathcal{N}, v) = \text{argmax}_{\mathcal{S}^c \in \Pi(S^c)} \sum_{T \in \mathcal{S}^c} v(T, \mathcal{S} \cup \mathcal{S}^c).$$

The following table summarizes the dependence of expectation formation rules on $\mathcal{S}$, $\mathcal{N}$ and v:

[3] We are aware that the argmax partition in this definition may not be unique and the same holds for the argmin in the pessimistic rule or the argmax in the max rule. Some tie-breaking rule can be used to choose a partition, but in order to avoid unnecessary notation, throughout the rest of the analysis we assume that the partition is unique up to symmetry considerations for the optimistic, pessimistic, and max rules.

rule	dep on $\mathcal{S}$	dep on $\mathcal{N}$	dep on v
Disintegration	no	yes	no
Projection	no	yes	no
$\mathcal{M}$-Exogenous	no	no	no
Optimistic	yes	no	yes
Pessimistic	yes	no	yes
Max	yes	no	yes

This table makes it clear that the expectation formation rules encountered in the literature can be organized into three kinds: Rules that depend on the original partition but not the reorganization of the deviating players or the partition function game, rules that depend on both the reorganization of the deviating players and the partition function game but not on the original partition, and one rule that does not depend on any of these three pieces of information. Perhaps this way of looking at expectation formation rules can be used to think about new rules that, for example, depend on $\mathcal{N}$ and v but not on $\mathcal{S}$ or rules that depend on all three pieces of information.

3 Axioms on expectation formation rules

In this section, we define axioms for expectation formation rules, and show how these axioms can be used to discriminate among different rules. We first introduce axioms on the dependence of $f(\cdot)$ with respect to the initial partition $\mathcal{N}$. We then present axioms relating expectations formed by players in a coalition and the expectations formed by players in smaller coalitions. We also introduce an axiom on the coherence of expectations formed by S and S^c. Finally, we discuss conditions under which an expectation formation rule preserves superadditivity by generating a superadditive coalitional function from any superadditive partition function.

3.1 Independence and responsiveness to $\mathcal{N}$

When external players react to the formation of $\mathcal{S}$ by coalition S, they can either be tied by the current partition, $\mathcal{N}$, or can freely reorganize independently of the original partition. In a more subtle way, the reaction of external players may or may not depend on the position in $\mathcal{N}$ of players perpetrating the deviation. The following axioms capture these different notions of independence.

Definition 3.1 *An expectation formation rule f is* independent of the original partition *(IOP) if $f(S, \mathcal{S}, \mathcal{N}, v) = f(S, \mathcal{S}, \mathcal{N}', v)$ for all $\mathcal{N}, \mathcal{N}' \in \Pi(N)$.*

Definition 3.2 *An expectation formation rule f is* independent of the position of deviating players in the original partition *(IPDOP) if $f(S, \mathcal{S}, \mathcal{N}, v) = f(S, \mathcal{S}, \mathcal{N}', v)$ for all $\mathcal{N}, \mathcal{N}' \in \Pi(N)$ such that $\mathcal{N}|_{S^c} = \mathcal{N}'|_{S^c}$.*

Definition 3.3 *An expectation formation rule f is* responsive to the position of external players in the original partition *(RPEOP) if $f(S, \mathcal{S}, \mathcal{N}, v) \neq f(S, \mathcal{S}, \mathcal{N}', v)$ for all $\mathcal{N}, \mathcal{N}' \in \Pi(N)$ such that $\mathcal{N}|_{S^c} \neq \mathcal{N}'|_{S^c}$.*

Notice that all usual rules but the disintegration rule and the projection rule are independent of the original partition. The disintegration rule does not satisfy IPDOP nor RPEOP, whereas the projection rule satisfies both axioms.

3.2 Path independence and subset consistency

The axioms of path dependence and subset consistency establish a connection between the expectations formed by different coalitions. Path independence states that, when a subset $S \cup T$ forms expectations, the expectations can either be formed first by S and then by T or first by T and then by S. In other words, the expectation formation rule must be independent of the order in which deviating agents form expectations. Subset consistency relates the expectations formed by a set S and any subset T of S and requires that these expectations be compatible, so that the projection of the expectations of members of T on S^c must be equal to the expectations of the members of S. The two axioms subset consistency and path independence guarantee consistency between expectations formed by different deviating coalitions and are formalized below.

Definition 3.4 *An expectation formation rule f satisfies* path independence *(PI) if, for any $S, T \subset N$ with $S, T \neq \emptyset$ and $S \cap T = \emptyset$, and for all $\mathcal{S} \in \Pi(S)$, $\mathcal{T} \in \Pi(T)$, $\mathcal{N} \in \Pi(N)$,*

$$f(S \cup T, \mathcal{S} \cup \mathcal{T}, \mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N}, v), v) = f(S \cup T, \mathcal{S} \cup \mathcal{T}, \mathcal{T} \cup f(T, \mathcal{T}, \mathcal{N}, v), v) \tag{2}$$

Definition 3.5 *An expectation formation rule f satisfies* subset consistency *(SC) if, for all $S \subseteq N, T \subset S$, $\mathcal{S} \in \Pi(S)$, $\mathcal{N} \in \Pi(N)$,*

$$f(T, \mathcal{S}|_T, \mathcal{N}, v)|_{S^c} = f(S, \mathcal{S}, \mathcal{N}, v). \tag{3}$$

Notice that, when subset consistency holds, the expectation formation rule cannot vary with the reorganization $\mathcal{S}$ of the deviating players.[4] Also, for an expectation formation rule that is subset consistent, RPEOP implies IPDOP, as we demonstrate in the following proposition.

Proposition 3.6 *If the expectations formation rule f satisfies subset consistency and is responsive to the position of external players in the original partition, then it is independent of the position of deviating players in the original partition.*

Proof: Because the expectation formation rule satisfies subset consistency, the partition $\mathcal{S}$ does not influence $f(S, \mathcal{S}, \mathcal{N}, v)$ and we omit $\mathcal{S}$ as an argument of the expectation formation rule f.
By RPEOP of f, we know that $f(S, \mathcal{N}, v) \neq f(S, \mathcal{N}', v)$ for all $\mathcal{N}, \mathcal{N}' \in \Pi(N)$ such that $\mathcal{N}|_{S^c} \neq \mathcal{N}'|_{S^c}$. From this we derive that

$$|\{f(S, \mathcal{N}, v) \mid \mathcal{N} \in \Pi(N)\}| \geq |\{\mathcal{N}|_{S^c} \mid \mathcal{N} \in \Pi(N)\}|.$$

Suppose that in addition to f satisfying RPEOP, there exist $\mathcal{N}', \mathcal{N}'' \in \Pi(N)$ with $\mathcal{N}'|_{S^c} = \mathcal{N}''|_{S^c}$ and $f(S, \mathcal{N}', v) \neq f(S, \mathcal{N}'', v)$. Then it follows that

$$|\{f(S, \mathcal{N}, v) \mid \mathcal{N} \in \Pi(N)\}| > |\{\mathcal{N}|_{S^c} \mid \mathcal{N} \in \Pi(N)\}|.$$

This, however, leads to a contradiction because

$$\{f(S, \mathcal{N}, v) \mid \mathcal{N} \in \Pi(N)\} \subseteq \Pi(S^c) = \{\mathcal{N}|_{S^c} \mid \mathcal{N} \in \Pi(N)\}.$$

We conclude that for all $\mathcal{N}', \mathcal{N}'' \in \Pi(N)$ with $\mathcal{N}'|_{S^c} = \mathcal{N}''|_{S^c}$ it must be the case that $f(S, \mathcal{N}', v) = f(S, \mathcal{N}'', v)$. Thus, f satisfies IPDOP. □

Path independence and subset consistency impose restrictions on cross-variations of the expectation formation rule on different variables: path independence considers variations in the original partition, whereas subset consistency focuses on variations in the set of deviating players. In spite of these differences, subset consistency implies path independence for expectation formation rules that are IPDOP.

Proposition 3.7 *If the expectation formation rule f satisfies subset consistency and independence of the position of deviating players in the original partition, then it satisfies path independence.*

[4]To see this, let $S \subset N$, $\mathcal{S}, \mathcal{S}' \in \Pi(S)$, and $i \in S$. Note that $\mathcal{S}|_{\{i\}} = \{i\} = \mathcal{S}'|_{\{i\}}$. Thus, it follows from subset consistency that $f(S, \mathcal{S}, v, \mathcal{N}) = f(i, \{i\}, v, \mathcal{N})|_{S^c} = f(S, \mathcal{S}', v, \mathcal{N})$. Note that Example 3.8 shows that there exist expectation formation rules that are independent of $\mathcal{S}$ but not subset consistent.

Proof: Let $\mathcal{N} \in \Pi(N)$. Consider two coalitions S, T such that $S \cap T = \emptyset$, and partitions $\mathcal{S} \in \Pi(S)$, $\mathcal{T} \in \Pi(T)$.
Applying subset consistency to coalitions S and $S \cup T$, we obtain

$$f(S, \mathcal{S}, \mathcal{N}, v)|_{(S \cup T)^c} = f(S \cup T, \mathcal{S} \cup \mathcal{T}, \mathcal{N}, v).$$

Because $\mathcal{S} \in \Pi(S)$, $f(S, \mathcal{S}, \mathcal{N}, v) \in \Pi(S^c)$, and $(S \cup T)^c \subseteq S^c$, adding $\mathcal{S}$ to $f(S, \mathcal{S}, \mathcal{N}, v)$ does not modify the projection onto $(S \cup T)^c$ so that

$$(\mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N}, v))|_{(S \cup T)^c} = f(S, \mathcal{S}, \mathcal{N}, v))|_{(S \cup T)^c}$$

Thus, we obtain

$$(\mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N}, v))|_{(S \cup T)^c} = f(S \cup T, \mathcal{S} \cup \mathcal{T}, \mathcal{N}, v). \tag{4}$$

Similarly, we derive

$$(\mathcal{T} \cup f(T, \mathcal{T}, \mathcal{N}, v))|_{(S \cup T)^c} = f(S \cup T, \mathcal{S} \cup \mathcal{T}, \mathcal{N}, v). \tag{5}$$

Given (4) and (5), we can apply IPDOP to obtain

$$f(S \cup T, \mathcal{S} \cup \mathcal{T}, \mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N}, v), v) = f(S \cup T, \mathcal{S} \cup \mathcal{T}, \mathcal{T} \cup f(T, \mathcal{T}, \mathcal{N}, v), v),$$

which demonstrates path independence. □

The following examples show that the two axioms of path independence and subset consistency are not equivalent.

Example 3.8 *(The expectation formation rule f satisfies path independence (and IOP) but not subset consistency)*
Let $N = \{i, j, k, l\}$. We define an expectation formation rule that only depends on the deviating coalitions S and let $f(S, \mathcal{S}, \mathcal{N}, v) = \underline{S^c}$ if $S = \{i\}$, and $f(S, \mathcal{S}, \mathcal{N}, v) = \overline{S^c}$ for all S such that $|S| \geq 2$.

This expectation formation rule obviously satisfies IOP (and thus also the weaker property IPDOP). The rule also satisfies path independence, because for all disjoint $S, T \subset N$ with $S, T \neq \emptyset$, we have $|S \cup T| \geq 2$ so that $f(S \cup T, \mathcal{S} \cup \mathcal{T}, \mathcal{N}, v) = \overline{(S \cup T)^c}$, independently of the partitions $\mathcal{S}$, $\mathcal{T}$, or $\mathcal{N}$.

The expectation formation rule does not satisfy subset consistency. To see this, note that $f(i, \{i\}, \mathcal{N}, v) = \{j|k|l\}$, so that $f(i, \{i\}, \mathcal{N}, v)|_{\{i,j\}^c} = \{k|l\}$, whereas $f(\{i, j\}, \{i, j\}, \mathcal{N}, v) = \{k, l\}$.[5]

[5] In examples, we use the less cluttered and commonly used notation of denoting a partition by separating the players in various coalitions with the symbol |. Hence, we write $\{j|kl\}$ instead of $\{\{j\}, \{k, l\}\}$ and so on.

The intuition for the discrepancy between path independence and subset consistency underlying Example 3.8 is that path independence does not impose any restrictions on the expectations of singletons, whereas subset consistency imposes a condition on the link between the expectations of singletons and those of larger coalitions. The following example illustrates that the requirement that f satisfies IPDOP cannot be omitted from the statement of Proposition 3.7.

Example 3.9 *(The expectation formation rule f satisfies subset consistency but not path independence)*
Suppose that $N = \{i, j, k, l\}$. We define an expectations rule that only depends on the deviating coalitions S and the partitions $\mathcal{N}$ and so we suppress $\mathcal{S}$ and v in the notation. Let $f(S, \mathcal{N}) = \underline{S^c}$ if $\mathcal{N} = \overline{N}$ or $\mathcal{N} = \{i|j|kl\}$, and $f(S, \mathcal{N}) = \mathcal{N}|_{S^c}$ otherwise.

This expectation formation rule satisfies subset consistency, because if $\mathcal{N} = \overline{N}$ or $\mathcal{N} = \{i|j|kl\}$, then $f(T, \mathcal{N})|_{S^c} = (\underline{T^c})\,|_{S^c} = \underline{S^c} = f(S, \mathcal{N})$, and for all other $\mathcal{N}$ it holds that $f(T, \mathcal{N})|_{S^c} = (\mathcal{N}|_{T^c})\,|_{S^c} = \mathcal{N}|_{S^c} = f(S, \mathcal{N})$. However, f violates path independence: Let $S = \{i\}$, $T = \{j\}$, and $\mathcal{N} = \{i|jkl\}$. When i forms expectations first, we obtain

$$
\begin{aligned}
f(S \cup T, \{S\} \cup f(S, \mathcal{N})) &= f(\{i, j\}, \{i\} \cup \{i|jkl\}|_{\{j,k,l\}}) \\
&= f(\{i, j\}, \{i|jkl\}) \\
&= \{i|jkl\}|_{\{k,l\}} = \{kl\}.
\end{aligned}
$$

However, when j forms expectations first, we obtain

$$
\begin{aligned}
f(S \cup T, \{T\} \cup f(T, \mathcal{N})) &= f(\{i, j\}, \{j\} \cup \{i|jkl\}|_{\{i,k,l\}}) \\
&= f(\{i, j\}, \{i|j|kl\}) \\
&= \underline{\{k, l\}} = \{k|l\}.
\end{aligned}
$$

3.3 Coherence of expectations

The next axiom imposes consistency between the formation of expectations of a coalition S and its complement S^c. Suppose that a coalition S contemplates reorganizing itself and forming a partition $\mathcal{S}$, expecting that the complement S^c reacts by forming $f(S, \mathcal{S}, \mathcal{N}, v)$. The axiom of coherence of expectations states that if indeed S^c forms this partition after S reorganizes and forms $\mathcal{S}$, members of S^c expect that the members of S will not subsequently reorganize again and form a partition different from $\mathcal{S}$.

Definition 3.10 *The expectation formation rule f satisfies* coherence of expectations *(COH) if, for all S, $\mathcal{S} \in \Pi(S)$ and $\mathcal{N} \in \Pi(N)$,*

$$f(S^c, f(S, \mathcal{S}, \mathcal{N}, v), \mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N}, v), v) = \mathcal{S}. \tag{6}$$

Coherence of expectations puts restrictions only on expectations held by a coalition and its complement, whereas subset consistency puts restrictions on expectations held by nested coalitions. Thus, the two axioms are independent, which is demonstrated in the next two examples.

Example 3.11 *(An $\mathcal{M}$-exogenous rule satisfies subset consistency and violates coherence of expectations.)*
Let $N = \{1, 2, 3, 4\}$, $\mathcal{M} = \{12|3|4\}$ and let f be the $\mathcal{M}$-exogenous rule. f satisfies subset consistency because $f(T, \mathcal{T}, \mathcal{N}, v)|_{S^c} = (\{12|3|4\}|_{T^c})\,|_{S^c} = \{12|3|4\}|_{S^c} = f(S, \mathcal{S}, \mathcal{N}, v)$ for all $T \subset S \subset N$. The expectations rule f does not satisfy coherence of expectations because, for example, for $S = \{1, 2\}$ and $\mathcal{S} = \{1|2\}$, it holds that $f(S^c, f(S, \mathcal{S}, \mathcal{N}, v), \mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N}, v), v) = \{12|3|4\}|_S = \{12\} \neq \{1|2\} = \mathcal{S}$.

Example 3.12 *(A rule that satisfies coherence of expectations but is not subset consistent.)*
Suppose that $N = \{i, j, k, l\}$. Define the expectations rule f that does not depend on v as follows. If $|S| = 1$ or $|S| = 3$, then $f(S, \mathcal{S}, \mathcal{N}) = \mathcal{N}|_{S^c}$. If $S = \{i, j\}$, then $f(S, \{i|j\}, \mathcal{N}) = \{k|l\}$ and $f(S, \{ij\}, \mathcal{N}) = \{kl\}$.

This rule satisfies coherence of expectations. This is seen as follows. If $|S| = 1$ or $|S| = 3$, then
$f(S^c, f(S, \mathcal{S}, \mathcal{N}), \mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N})) = f(S^c, \mathcal{N}|_{S^c}, \mathcal{S} \cup \mathcal{N}|_{S^c}) = (\mathcal{S} \cup \mathcal{N}|_{S^c})|_S = \mathcal{S}$.
If $S = \{i, j\}$, and $\mathcal{S} = \{i|j\}$, then
$f(S^c, f(S, \mathcal{S}, \mathcal{N}), \mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N})) = f(\{k, l\}, \{k|l\}, \{i|j\} \cup \{k|l\}) = \{i|j\} = \mathcal{S}$.
If $S = \{i, j\}$ and $\mathcal{S} = \{ij\}$, then
$f(S^c, f(S, \mathcal{S}, \mathcal{N}), \mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N})) = f(\{k, l\}, \{kl\}, \{ij\} \cup \{kl\}) = \{ij\} = \mathcal{S}$.

The rule violates subset consistency, because with $\mathcal{N} = \{1234\}$, $S = \{1, 2\}$, $\mathcal{S} = \{1|2\}$, and $T = \{1\}$, we have that $f(T, \mathcal{S}|_T, \mathcal{N})|_{S^c} = \{34\} \neq \{3|4\} = f(S, \mathcal{S}, \mathcal{N})$.

3.4 Preservation of superadditivity

The next axiom pertains to the superadditivity of the coalitional functions $w_f^{\mathcal{N}}$ generated by the expectation formation rule f.

Definition 3.13 *The expectation formation rule f satisfies* preservation of superadditivity *(PSA) if, whenever v is a superadditive partition function, $w_f^{\mathcal{N}}$ is a superadditive coalitional function for all $\mathcal{N} \in \Pi(N)$.*

The preservation of superadditivity is not obvious because in the partition function game the partition of external players is kept constant when two coalitions merge, whereas in the coalitional game the external players may form different partitions depending on the organization of the merging players.

4 Axiomatizations of expectation formation rules

In this section we demonstrate that the axioms on expectation formation rules that we identified in the previous section can be used to axiomatize some of the rules. We first consider rules that satisfy responsiveness to the position of external players in the original partition and find that the projection rule takes a special position in the class of responsive expectation formation rules. We then consider rules that are independent of the original partition and find that the $\mathcal{M}$-exogenous rules and the pessimistic rule are the most prominent independent expectation formation rules from an axiomatic point of view.

4.1 Responsive rules

We first axiomatize rules which depend on the current partition $\mathcal{N}$. The following theorem demonstrates that the projection rule is the only subset consistent rule among the responsive expectation formation rules.

Theorem 4.1 *Let $n \geq 4$. An expectation formation rule f satisfies subset consistency and responsiveness to the position of external players in the original partitions if and only if it is the projection rule.*

Proof: It is clear that the projection rule satisfies subset consistency and RPEOP. Now, consider an expectation formation rule f that satisfies the two axioms. Because the expectation formation rule satisfies subset consistency, the partition $\mathcal{S}$ does not influence $f(S, \mathcal{S}, \mathcal{N}, v)$ and we omit $\mathcal{S}$ as an argument of the expectation formation rule f. Also, by Proposition 3.6 f satisfies IPDOP and thus for any $\mathcal{N} \in \Pi(N)$ it holds that

$$f(S, \mathcal{N}, v) = f(S, \mathcal{U} \cup \mathcal{N}|_{S^c}, v) \text{ for any } \mathcal{U} \in \Pi(S). \tag{7}$$

Using this, and with minimal abuse of notation, we can write $f(S, \mathcal{N}|_{S^c}, v)$ whenever we do not want to explicitly specify the behavior of $\mathcal{N}$ on players not in S^c.

First notice that if $|S| = n-1$, then $S^c = \{i\}$ for some $i \in N$ and trivially $f(S, \mathcal{N}, v) = \{i\}$ for all $\mathcal{N}$.

Claim 4.2 *For any $S \subseteq N$ such that $|S| = n-2$, and any $\mathcal{N} \in \Pi(N)$ it holds that $f(S, \mathcal{N}, v) = \mathcal{N}|_{S^c}$.*

Proof of the Claim: Because the expectations rule f satisfies RPEOP, it must assign a different partition to every $\mathcal{N}|_{S^c} \in \Pi(S^c)$ when taken as given a coalition $S \subset N$ and partition function game v. Thus, for every pair of players $i, j \in N$, either

$$f(\{i,j\}^c, \{ij\}, v) = \{ij\} \text{ and } f(\{i,j\}^c, \{i|j\}, v) = \{i|j\}$$

or

$$f(\{i,j\}^c, \{ij\}, v) = \{i|j\} \text{ and } f(\{i,j\}^c, \{i|j\}, v) = \{ij\}.$$

Hence, once we determine $f(\{i,j\}^c, \{ij\}, v)$, we have no flexibility in choosing the expectation $f(\{i,j\}^c, \{i|j\}, v)$.

Consider three players, 1, 2 and 3, the set $T = \{1,2,3\}^c$, and the three sets $S_1 = \{2,3\}^c$, $S_2 = \{1,3\}^c$ and $S_3 = \{1,2\}^c$. Notice that $T \neq \emptyset$ as $n \geq 4$. In what follows, we let $\{i,j,k\} = \{1,2,3\}$. Given that the expectation formation rule satisfies RPEOP, it is sufficient to construct $f(S_i, \{jk\}, v)$ for each S_i, so there are eight ways in which we can construct the partitions $f(S_i, \mathcal{N}|_{S_i^c}, v)$, $i = 1,2,3$. Disregarding cases which are symmetric up to a permutation of the players, we only need to consider four different cases: (i) the case where $f(S_i, \{jk\}, v) = \{jk\}$ for all $i = 1,2,3$, (ii) the case where $f(S_i, \{jk\}, v) = \{jk\}$ for two players $i \in \{1,2,3\}$, and $f(S_i, \{jk\}, v) = \{j|k\}$ for the third player, (iii) the case where $f(S_i, \{jk\}, v) = \{jk\}$ for one player $i \in \{1,2,3\}$, and $f(S_i, \{jk\}, v) = \{j|k\}$ for the other two players and (iv) the case where $f(S_i, \{jk\}, v) = \{j|k\}$ for all three players.

Now consider the expectations of players in $T = \{1,2,3\}^c$. T is a subset of S_i for each $i \in \{1,2,3\}$. We will use subset consistency to prove that cases (ii), (iii) and (iv) result in a contradiction.

Consider case (ii) when $f(S_1, \{23\}, v) = \{23\}$, $f(S_2, \{13\}, v) = \{13\}$, and $f(S_3, \{12\}, v) = \{1|2\}$. Then, by subset consistency, for a partition $\mathcal{N} \in \Pi(N)$ such that $\mathcal{N}|_{T^c} = \{123\}$,

$$
\begin{aligned}
f(T,\{123\},v)|_{\{12\}} &= f(S_3,\{12\},v) &= \{1|2\}\\
f(T,\{123\},v)|_{\{13\}} &= f(S_2,\{13\},v) &= \{13\}\\
f(T,\{123\},v)|_{\{23\}} &= f(S_1,\{23\},v) &= \{23\}
\end{aligned}
$$

resulting in a contradiction, as we cannot find a partition $f(T,\{123\},v)$ of $\{123\}$ that projects into $\{1|2\}$, $\{13\}$, and $\{23\}$.

Consider case (iii) when $f(S_1,\{23\},v) = \{23\}$, $f(S_2,\{13\},v) = \{1|3\}$, and $f(S_3,\{12\},v) = \{1|2\}$. Again, by subset consistency,

$$
\begin{aligned}
f(T,\{1|2|3\},v)|_{\{12\}} &= f(S_3,\{1|2\},v) &= \{12\}\\
f(T,\{1|2|3\},v)|_{\{13\}} &= f(S_2,\{1|3\},v) &= \{13\}\\
f(T,\{1|2|3\},v)|_{\{23\}} &= f(S_1,\{2|3\},v) &= \{2|3\}
\end{aligned}
$$

resulting in a contradiction because we cannot find a partition $f(T,\{1|2|3\},v)$ of $\{123\}$ that projects into $\{12\}$, $\{13\}$, and $\{2|3\}$.

Finally, in case (iv), consider

$$
\begin{aligned}
f(T,\{1|23\},v)|_{\{12\}} &= f(S_3,\{1|2\},v) &= \{12\}\\
f(T,\{1|23\},v)|_{\{13\}} &= f(S_2,\{1|3\},v) &= \{13\}\\
f(T,\{1|23\},v)|_{\{23\}} &= f(S_1,\{23\},v) &= \{2|3\}
\end{aligned}
$$

resulting in a contradiction because we cannot find a partition $f(T,\{1|23\},v)$ of $\{123\}$ that projects into $\{12\}$, $\{13\}$, and $\{2|3\}$.

Since cases (ii), (iii), and (iv) all lead to a contradiction, we are left the conclusion that case (i) must hold, which proves the claim.

We finish the proof of the theorem by induction. Let $m < n$ such that $m \geq 3$ and suppose that we have shown that $f(S,\mathcal{N}|_{S^c},v) = \mathcal{N}|_{S^c}$ for all coalitions S such that $|S^c| < m$, and any $\mathcal{N} \in \Pi(N)$. Consider a set T such that $|T^c| = m$. For all $i \in T^c$, define the set $S_i := T \cup i$. Let $\mathcal{N} \in \Pi(N)$. For each $i \in T^c$, we have $T \subset S_i$ and $|S_i^c| = m-1$, and thus by applying subset consistency and the induction hypothesis, we obtain

$$
f(T,\mathcal{N}|_{T^c},v)|_{S_i^c} = f(S_i,\mathcal{N}|_{S_i^c},v) = \mathcal{N}|_{S_i^c}. \tag{8}
$$

Fix two players $i, j \in T^c$. Then either i and j belong to different blocks in the partition $\mathcal{N}$ or they belong to the same block in the partition $\mathcal{N}$. Because $|T^c| = m \geq 3$, we can find a player $k \in T^c$, $k \notin \{i, j\}$, and by equation (8) we know that for the set $S_k = T \cup k$

$$f(T, \mathcal{N}|_{T^c}, v)|_{S_k^c} = \mathcal{N}|_{S_k^c}.$$

Note that i and j do not belong to S_k. It thus follows that i and j belong to different blocks in the partition $f(T, \mathcal{N}|_{T^c}, v)$ if and only if they belong to different blocks in the partition $\mathcal{N}|_{S_k^c}$, and they belong to different blocks in the partition $\mathcal{N}|_{S_k^c}$ if and only if they belong to different blocks in $\mathcal{N}$. This establishes that $f(T, \mathcal{N}|_{T^c}, v) = \mathcal{N}|_{T^c}$, completing the proof of the theorem. □

Theorem 4.1 characterizes the projection rule as the only responsive rule that satisfies subset consistency. Notice that this characterization only holds for $n \geq 4$. For $n = 3$, we can find responsive and subset consistent rules that are not the projection rule, as is shown in the following example.

Example 4.3 *(A rule that is responsive and subset consistent for $n = 3$ and that does not coincide with the projection rule).*
Suppose that $N = \{i, j, k\}$. Define the expectation formation rule f as follows. If $S = \{i, j\}$, then $f(S, \mathcal{S}, \mathcal{N}, v) = \{k\}$. If $S = \{i\}$ and $\mathcal{N}|_{S^c} = \{jk\}$, then $f(S, \{i\}, \mathcal{N}, v) = \{j|k\}$. If $S = \{i\}$ and $\mathcal{N}|_{S^c} = \{j|k\}$, then $f(S, \{i\}, \mathcal{N}, v) = \{jk\}$.

Clearly, f satisfies RPEOP. It also satisfies subset consistency, because the only possible choices for two nested coalitions $T \subset S \subseteq N$ are $S = \{i, j\}$ and $T = \{i\}$ and then $f(T, \mathcal{S}|_T, \mathcal{N}, v)|_{S^c} = \{k\}$ because the only possible partition of a singleton is a singleton. However, the rule f is not the projection rule.

An alternative characterization of the projection rule can be given in terms of subset consistency and coherence of expectations.

Theorem 4.4 *An expectation formation rule f satisfies subset consistency, independence of the position of deviating players in the original partition and coherence of expectations if and only if it is the projection rule.*

Proof: It is easy to check that the projection rule satisfies coherence of expectations and IPDOP in addition to subset consistency. Now, consider an expectation formation rule f that satisfies the three axioms. Because the expectation formation rule satisfies subset consistency, $\mathcal{S}$ does not influence[6] $f(S,\mathcal{S},\mathcal{N},v)$ and because the expectation formation rule satisfies IPDOP, we know that the behavior of $\mathcal{N}$ on S does not influence $f(S,\mathcal{S},\mathcal{N},v)$. Let $S \subseteq N$, $\mathcal{S} \in \Pi(S)$, and $\mathcal{N} \in \Pi(N)$. We obtain

$$\begin{aligned} f(S,\mathcal{S},\mathcal{N},v) &= f(S, f(S^c,\mathcal{N}|_{S^c},\mathcal{N},v),\mathcal{N},v) \\ &= f(S, f(S^c,\mathcal{N}|_{S^c},\mathcal{N},v),\mathcal{N}|_{S^c} \cup f(S^c,\mathcal{N}|_{S^c},\mathcal{N},v),v) \\ &= \mathcal{N}|_{S^c}, \end{aligned}$$

where the first equality follows from subset consistency (changing the partition of S from $\mathcal{S}$ to $f(S^c,\mathcal{N}|_{S^c},\mathcal{N},v)$ has no influence on the expectation), the second equality follows from IPDOP because $(\mathcal{N}|_{S^c} \cup f(S^c,\mathcal{N}|_{S^c},\mathcal{N},v))\,|_{S^c} = \mathcal{N}|_{S^c}$, and the third equality follows by applying coherence of expectations (with the roles of S^c and S interchanged). This shows that f is the projection rule. $\square$

The three axioms in Theorem 4.4 are logically independent. The rule in Example 3.11 satisfies subset consistency and IPDOP, but violates coherence of expectations. The rule in Example 3.12 satisfies coherence of expectations and IPDOP, but is not subset consistent. Finally, the next example displays a rule that satisfies subset consistency and coherence of expectations, but violates IPDOP.

Example 4.5 *(A rule that satisfies subset consistency and coherence of expectations, but violates independence of the position of deviating players in the original partition.)*
We define an expectation formation rule f that does not depend on $\mathcal{S}$ or v and we simplify notation accordingly. We define $f(S,\mathcal{N}) = \mathcal{N}|_{S^c}$ if $\mathcal{N} \neq \{N\}$ and $f(S,\{N\}) = \underline{S^c}$.

f satisfies subset consistency because for any $T \subset S \subseteq N$ it holds that $f(T,\mathcal{N})|_{S^c} = (\mathcal{N}|_{T^c})\,|_{S^c} = \mathcal{N}|_{S^c} = f(S,\mathcal{N})$ if $\mathcal{N} \neq \{N\}$, while $f(T,\{N\})|_{S^c} = \underline{T^c}|_{S^c} = \underline{S^c} = f(S,\mathcal{N})$.

f satisfies coherence of expectations because $f(S^c,\mathcal{S} \cup f(S,\mathcal{N})) = f(S^c,\mathcal{S} \cup \mathcal{N}|_{S^c}) = (\mathcal{S} \cup \mathcal{N}|_{S^c})|_S = \mathcal{S}$ if $\mathcal{N} \neq \{N\}$, while $f(S^c,\mathcal{S} \cup f(S,\{N\})) = f(S^c,\mathcal{S} \cup \underline{S^c}) = (\mathcal{S} \cup \underline{S^c})|_S = \mathcal{S}$.

f does not satisfy IPDOP and indeed is not the projection rule.

[6] See footnote 4.

4.2 Independent expectation formation rules

In this subsection, we consider expectation formation rules that do not depend on the current partition $\mathcal{N}$. Our first result points out that subset consistency then results in exogenous projections.

Theorem 4.6 *An expectation formation rule f satisfies subset consistency and independence of the original partition if and only if it is an exogenous rule.*

Proof: It is clear that for any $\mathcal{M}$, the $\mathcal{M}$-exogenous rule satisfies subset consistency and independence of the original partition. Now, consider an expectation formation rule f that satisfies these two axioms. This implies that neither the new partition of deviating players $\mathcal{S}$ nor the original partition $\mathcal{N}$ influence the expectations, and the expectation formation rule only depends on S and v. To economize notation, we let $f(S, v)$ denote the expectation formation rule throughout the remainder of this proof.

To prove that f is an exogenous rule, we need to show that for any two players $i, j \in N$ and any two coalitions $S_1, S_2 \subseteq \{i, j\}^c$, it holds that i and j are in the same block in the partition $f(S_1, v)$ if and only if they are in the same block in the partition $f(S_2, v)$ or, equivalently, that $f(S_1, v)|_{\{i,j\}} = f(S_2, v)|_{\{i,j\}}$. But this follows directly from subset consistency, which implies

$$f(S_1, v)|_{(S_1 \cup S_2)^c} = f(S_1 \cup S_2, v) = f(S_2, v)|_{(S_1 \cup S_2)^c}.$$

Notice that $\{i, j\} \subseteq (S_1 \cup S_2)^c$, so that i and j belong to the same block in $f(S_1, v)$ if and only if they belong to the same block in $f(S_2, v)$. □

Theorem 4.6 implicitly points out that common independent expectation formation rules such as the optimistic, pessimistic and max expectation rules, do not satisfy subset consistency and result in an inconsistency in the expectation of a coalition of deviating players and a subset of this coalition.

We now turn to preservation of superadditivity. For an expectation formation rule f that is independent of the original partition, the coalitional function $w_f^{\mathcal{N}}$ is the same for all $\mathcal{N}$ and thus there is a unique coalitional function that is generated by f and we refer to this function as w_f. We show in the next proposition that when the expectation formation rule is the pessimistic rule, then the coalitional game w_f is superadditive.

Proposition 4.7 *The pessimistic rule satisfies preservation of superadditivity.*

Proof: Let f be the pessimistic rule. We simplify notation by suppressing the original partition and write $f(S, \mathcal{S}, v)$.

Let $S, T \subset N$, $S, T \neq \emptyset$, with $S \cap T = \emptyset$. Define

$$\hat{\mathcal{S}} = \arg \max_{\mathcal{S} \in \Pi(S)} \sum_{S_i \in \mathcal{S}} v(S_i, \mathcal{S} \cup f(S, \mathcal{S}, v))$$

and let $\hat{\mathcal{T}}$ be defined similarly. The partition $\hat{\mathcal{S}}$ ($\hat{\mathcal{T}}$) is the partition that gives S (T) the maximal worth given its expectations according to the pessimistic expectation formation rule f and thus

$$w_f(S) = \sum_{S_i \in \hat{\mathcal{S}}} v(S_i, \hat{\mathcal{S}} \cup f(S, \hat{\mathcal{S}}, v))$$

and

$$w_f(T) = \sum_{T_i \in \hat{\mathcal{T}}} v(T_i, \hat{\mathcal{T}} \cup f(T, \hat{\mathcal{T}}, v)).$$

The worth $w_f(S \cup T)$ is obtained when the members of $S \cup T$ organize themselves into a partition that maximizes their worth, expecting that the other players will form a partition that minimizes their worth of the players $S \cup T$. Since $\hat{\mathcal{S}} \cup \hat{\mathcal{T}}$ is a partition of $S \cup T$ that may or may not be optimal for $S \cup T$,

$$\begin{aligned} w_f(S \cup T) \;\; \geq \;\; & \sum_{S_i \in \hat{\mathcal{S}}} v(S_i, \hat{\mathcal{S}} \cup \hat{\mathcal{T}} \cup f(S \cup T, \hat{\mathcal{S}} \cup \hat{\mathcal{T}}, v)) \\ & + \sum_{T_i \in \hat{\mathcal{T}}} v(T_i, \hat{\mathcal{S}} \cup \hat{\mathcal{T}} \cup f(S \cup T, \hat{\mathcal{S}} \cup \hat{\mathcal{T}}, v)). \end{aligned}$$

Because the expectations $f(S, \hat{\mathcal{S}}, v)$ are pessimistic,

$$\begin{aligned} \sum_{S_i \in \hat{\mathcal{S}}} v(S_i, \hat{\mathcal{S}} \cup \hat{\mathcal{T}} \cup f(S \cup T, \hat{\mathcal{S}} \cup \hat{\mathcal{T}}, v)) \;\; \geq \;\; & \sum_{S_i \in \hat{\mathcal{S}}} v(S_i, \hat{\mathcal{S}} \cup f(S, \hat{\mathcal{S}}, v)) \\ & = w_f(S). \end{aligned}$$

Similarly,

$$\sum_{T_i \in \hat{\mathcal{T}}} v(T_i, \hat{\mathcal{S}} \cup \hat{\mathcal{T}} \cup f(S \cup T, \hat{\mathcal{S}} \cup \hat{\mathcal{T}}, v)) \geq w_f(T),$$

so that $w_f(S \cup T) \geq w_f(S) + w_f(T)$ follows. □

It is interesting to note that the coalitional function generated by the pessimistic rule is superadditive even when the underlying partition function game fails superadditivity. As the following example shows, preservation of superadditivity is a very strong requirement, and other commonly used expectation rules that are independent of the original partition fail to satisfy this axiom.

Example 4.8 *(A superadditive partition function game that does not generate a superadditive coalitional game for usual independent expectation formation rules other than the pessimistic rule.)*
Let $N = \{i, j, k, l\}$. *Consider the superadditive symmetric partition function game* v *defined by* $v(i, i|j|k|l) = 2$, $v(ij, ij|k|l) = 7$, $v(k, ij|k|l) = 0$, $v(ij, ij|kl) = 10$, $v(ijk, ijk|l) = 8$, $v(l, ijk|l) = 4$, $v(ijkl, ijkl) = 21$.

Note that $w_f(N)$ *is independent of the expectation formation rule* f. *In this example, it is accomplished in the partition* $\overline{N}$ *and equals 21. Also, when* $|S| = n-1$, *then* $S^c = \{i\}$ *for some* $i \in N$ *and necessarily* $f(S, \mathcal{S}, \mathcal{N}, v) = \{i\}$, *no matter how the rule* f *is defined. Thus,* $w_f(S)$ *is independent of the expectation formation rule that is used, and in this example it is equal to 8.*

If the expectation formation rule f *is the* $\underline{N}$*-exogenous rule, then we derive* $w_f(i) = v(i, i \cup j|k|l) = 2$ *and* $w_f(i, j) = \max\{v(ij, ij \cup k|l), v(i, i|j \cup k|l) + v(j, i|j \cup k|l)\} = 7$. *This coalitional game is shown in the second column of the table below.*

We list the values for the coalitional games w_f *for other IOP expectation formation rules* f *without computations.*

$\lvert S\rvert$	$\underline{N}$-*exogenous*	$\overline{N}$-*exogenous*	*optimistic*	*pessimistic*	*max*
1	2	4	4	0	4
2	7	10	10	7	10
3	8	8	8	8	8
4	21	21	21	21	21

While the coalition game w_f *obtained when* f *is the pessimistic rule is superadditive, all the coalitional games derived from the other IOP rules do not satisfy superadditivity because for the games* w_f *in the other columns it holds that* $w_f(i) + w_f(j, k) > w_f(i, j, k)$.

4.3 Summary of properties of expectation formation rules

The table below summarizes the properties satisfied by the usual expectation formation rules.

	IOP	IPDOP	RPEOP	PI	SC	COH	PSA
Disintegration				✓		✓	
Projection		✓	✓	✓	✓	✓	
$\mathcal{M}$-Exogenous	✓	✓		✓	✓		
Optimistic	✓	✓		✓			
Pessimistic	✓	✓		✓			✓
Max	✓	✓		✓			

Most verifications have been covered in the preceding subsections or are immediate. The remaining (lack of) checkmarks in the table are addressed in an appendix.

An interesting observation from the table is that path independence does not allow us to distinguish between the various commonly used expectation formation rules. The other consistency axioms - subset consistency and coherence of expectations - are much more discriminating.

5 Cores of partition function games

Each expectation formation rule gives rise to a different definition of the core of a partition function game. The optimistic core is the smallest and the pessimistic one the largest. In this section we also extend balancedness to partition function games. We conclude this section by highlighting the core based on the projection rule, which is singled out because of the summary table in Subsection 4.3.

5.1 Expectation formation rules and cores of partition function games

Given an expectation formation rule f and a partition $\mathcal{N}$, we construct the TU coalitional game $w_f^{\mathcal{N}}$ as in equation (1):

$$w_f^{\mathcal{N}}(S) = \max_{\mathcal{S}\in\Pi(S)} \sum_{T\in\mathcal{S}} v(T, \mathcal{S}\cup f(S,\mathcal{S},\mathcal{N},v)).$$

The cores of these games - the set of imputations that are immune to deviations by any coalition - will obviously depend on the expectation formation rule used and also on the original partition. Because the partition function game v is not necessarily superadditive, we do not assume that the grand coalition $\overline{N}$ forms. Instead, in order to prevent deviations by the grand coalition, we use the partition that maximizes the sum of payoffs of all players:

$$\mathcal{N}^* = \text{argmax}_{\mathcal{N}\in\Pi(N)} \sum_{S\in\mathcal{N}} v(S,\mathcal{N}).$$

Definition 5.1 *The core $C_f(v)$ of partition function game v with respect to expectation formation rule f is the set of vectors $(x_1, x_2, ..., x_n)$ in $\Re^n$ that satisfy the following two conditions:*

1. $\sum_{i\in N} x_i = \sum_{S\in\mathcal{N}^*} v(S,\mathcal{N}^*)$
2. $\sum_{i\in S} x_i \geq w_f^{\mathcal{N}^*}(S)$ *for all coalitions* $S \subseteq N$.

Keeping the partition function game v fixed and denoting the optimistic and pessimistic expectation formation rules by o and p, respectively, we have that for any expectation formation rule f and all $S \subseteq N$,

$$w_p(S) \leq w_f^{\mathcal{N}^*}(S) \leq w_o(S).$$

Thus, the optimistic core the smallest core and the the pessimistic core is the largest core:

$$C_o(v) \subseteq C_f(v) \subseteq C_p(v)$$

for all expectation formation rules f.

5.2 Balancedness of partition function games

One way of trying to select a core generated by a particular expectation formation rule is to parallel the balancedness approach for coalitional games. The core of a coalitional game is a convex polytope characterized by a set of linear inequalities. In order to guarantee the existence of a solution to the set of inequalities, one can consider the dual linear programming problem, resulting in the definition of balanced coalitional games.

By following a similar approach for partition function games, we are led to define the set of *embedded coalitions* $E(N) = \{(S,\mathcal{N}) \mid S \in \mathcal{N} \in \Pi(N)\}$ and weights $\delta(S,\mathcal{N}) \geq 0$, $(S,\mathcal{N}) \in E(N)$. A collection of embedded coalitions $\mathcal{E} \subseteq E(N)$ is balanced if there exist *balancing weights* $(\delta(S,\mathcal{N}))_{(S,\mathcal{N})\in\mathcal{E}}$ such that for each $i \in N$

$$\sum_{(S,\mathcal{N})\in\mathcal{E}:\, i\in S} \delta(S,\mathcal{N}) = 1.$$

Definition 5.2 *A partition function game v is* balanced *if, for any balanced collection of embedded coalitions $\mathcal{E}$ with balancing weights $(\delta(S,\mathcal{N}))_{(S,\mathcal{N})\in\mathcal{E}}$,*

$$\sum_{(S,\mathcal{N})\in\mathcal{E}:\, i\in S} \delta(S,\mathcal{N})\, v(S,\mathcal{N}) \leq \sum_{S\in\mathcal{N}^*} v(S,\mathcal{N}^*).$$

The extension of balancedness to partition function games is related to the optimistic expectation formation rule, as evidenced by the next proposition.

Proposition 5.3 *A partition function game v is balanced if and only if its optimistic core $C_o(v)$ is nonempty.*

Proof: Let v be a partition function form game. Consider the linear program

$$\begin{aligned} \text{maximize} \quad & \sum_{(S,\mathcal{N})\in\mathcal{E}:\, i\in S} \delta(S,\mathcal{N})\, v(S,\mathcal{N}) \\ \text{subject to} \quad & \sum_{(S,\mathcal{N})\in\mathcal{E}:\, i\in S} \delta(S,\mathcal{N}) = 1 \text{ for each } i\in N \\ & \delta(S,\mathcal{N}) \geq 0 \text{ for all } (S,\mathcal{N})\in E(N) \end{aligned} \tag{9}$$

and its dual

$$\begin{aligned} \text{minimize} \quad & \sum_{i\in N} x_i \\ \text{subject to} \quad & \sum_{i\in S} x_i \geq v(S,\mathcal{N}) \text{ for all } (S,\mathcal{N})\in E(N) \end{aligned} \tag{10}$$

The duality theorem tells us that the optimal values of the two programs are the same if they have a solution. We observe that the optimal value of (9) is at least $\sum_{S\in\mathcal{N}^*} v(S,\mathcal{N})$, because this value is attained for the balancing weights $\delta(S,\mathcal{N}^*) = 1$ for each $S\in\mathcal{N}^*$ and $\delta(S,\mathcal{N}) = 0$ for all other embedded coalitions.

It follows from the definitions that v is balanced if and only if the optimal value of (9) equals $\sum_{S\in\mathcal{N}^*} v(S,\mathcal{N})$. By the duality theorem this is the case if and only if the optimal value of (10) equals $\sum_{S\in\mathcal{N}^*} v(S,\mathcal{N})$. In turn, this is the case if and only if there exist $(x_1, x_2, ..., x_n)\in\Re^n$ such that $\sum_{i\in S} x_i \geq v(S,\mathcal{N})$ for all $(S,\mathcal{N})\in E(N)$ and $\sum_{i\in N} x_i = \sum_{S\in\mathcal{N}^*} v(S,\mathcal{N})$. The latter is the case if and only if $C_o(v)\neq\emptyset$ (because $v(S,\mathcal{N}) \leq w_o(S)$ for all $S\subseteq N$ and all $\mathcal{N}\in\Pi(N)$). □

Proposition 5.3 shows that balancedness of the partition function game is equivalent to the optimistic core of the game being nonempty. However, because the optimistic core is the smallest core, for a balanced game the cores generated by all expectation formation rules are nonempty.

5.3 The projection core of superadditive partition function games

Based on the axiomatizations in Section 4, an important intermediate core is the *projection core* $C_{pr}(v)$, which is the core of the coalitional game $w_{pr}^{\mathcal{N}^*}$ generated by the projection rule pr. This core is easy to compute, especially for superadditive partition function games with positive externalities.

Proposition 5.4 *In a superadditive partition function game with positive externalities, $\mathcal{N}^* = \overline{N}$, and $w_{pr}^{\mathcal{N}^*}(S) = v(S, \overline{S} \cup \overline{S^c})$. Moreover, for such games the the projection core coincides with the optimistic core.*

Proof: Let v be a superadditive partition function form game with positive externalities. If $\mathcal{N}$ is a partition of N that has a at least two blocks $S, T \in \mathcal{N}$, then the sum of payoffs of all players does not decrease when the blocks S and T join:

$$\begin{aligned} v(S \cup T, \mathcal{N} \setminus \{S,T\} \cup \{S \cup T\}) &+ \sum_{U \in \mathcal{N} \setminus \{S,T\}} v(U, \mathcal{N} \setminus \{S,T\} \cup \{S \cup T\}) \\ &\geq v(S,\mathcal{N}) + v(T,\mathcal{N}) + \sum_{U \in \mathcal{N} \setminus \{S,T\}} v(U, \mathcal{N} \setminus \{S,T\} \cup \{S \cup T\}) \\ &\geq v(S,\mathcal{N}) + v(T,\mathcal{N}) + \sum_{U \in \mathcal{N} \setminus \{S,T\}} v(U,\mathcal{N}) \\ &= \sum_{U \in \mathcal{N}} v(U,\mathcal{N}), \end{aligned}$$

where the first inequality follows from superaditivity, and the second one from positive externalities. Thus, $\mathcal{N}^* = \overline{N}$.

By definition of the projection rule, $pr(S, \mathcal{S}, \overline{N}, v) = \overline{S^c}$ for all S and $\mathcal{S}$. If $\mathcal{S}$ is a partition of S that has a at least two blocks $S_1, S_2 \in \mathcal{S}$, then with expectations according to the projection rule the sum of payoffs of all the players in S does not decrease when the blocks S_1 and S_2 join:

$$\begin{aligned}
v(S_1 \cup S_2, \mathcal{S} \setminus \{S_1, S_2\} \cup \{S_1 \cup S_2\} \cup \overline{S^c}) &+ \\
&\sum_{U \in \mathcal{S} \setminus \{S_1, S_2\}} v(U, \mathcal{S} \setminus \{S_1, S_2\} \cup \{S_1 \cup S_2\} \cup \overline{S^c}) \\
&\geq v(S_1, \mathcal{S} \cup \overline{S^c}) + v(S_2, \mathcal{S} \cup \overline{S^c}) + \\
&\sum_{U \in \mathcal{S} \setminus \{S_1, S_2\}} v(U, \mathcal{S} \setminus \{S_1, S_2\} \cup \{S_1 \cup S_2\} \cup \overline{S^c}) \\
&\geq v(S_1, \mathcal{S} \cup \overline{S^c}) + v(S_2, \mathcal{S} \cup \overline{S^c}) + \sum_{U \in \mathcal{S} \setminus \{S_1, S_2\}} v(U, \mathcal{S} \cup \overline{S^c}) \\
&= \sum_{U \in \mathcal{S}} v(U, \mathcal{S} \cup \overline{S^c}),
\end{aligned}$$

where the first inequality follows from superaditivity, and the second one from positive externalities. From this it follows that

$$\text{argmax}_{\mathcal{S} \in \Pi(S)} \sum_{T \in \mathcal{S}} v(T, \mathcal{S} \cup \overline{S^c}) = \overline{S},$$

so that

$$w_{pr}^{\overline{N}}(S) \;\; = \;\; v(S, \overline{S} \cup \overline{S^c}),$$

the worth of S in the two-block partition $\mathcal{N} = \{S, S^c\}$ consisting of S and its complement S^c.

Because v has positive externalities, the optimistic expectations of each coalition S are $o(S, \mathcal{S}, \overline{N}, v) = \overline{S^c}$, so that they coincide with the expectations according to the projection rule when the original partition is $\mathcal{N}^* = \overline{N}$. From this, we easily derive that

$$w_o(S) = w_{pr}^{\overline{N}}(S) = v(S, \overline{S} \cup \overline{S^c})$$

for all $S \subseteq N$, so that both of these coalitional games have the same core. Thus, the projection core and the optimistic core coincide. □

One might suspect that Proposition 5.4 has a counterpart for superadditive partition function games with negative externalities and that for such games the projection core coincides with the pessimistic core. Such a result, however, does not hold. This is illustrated in the following example, which also illustrates that superadditivity of a partition function game is not sufficient for the grand coalition to be payoff-maximizing for N.

Example 5.5 *Let $N = \{i, j, k, l\}$. Consider a symmetric partition function game where $v(i, i|j|k|l) = 4$, $v(ij, ij|k|l) = 9$, $v(k, ij|k|l) = 1$, $v(ij, ij|kl) = 3$, $v(ijk, ijk|l) = 11$, $v(l, ijk|l) = 0$, $v(ijkl, ijkl) = 12$. This partition function game satisfies superadditivity and has negative externalities.*

Even though the partition function game v is superadditive, the coalition structure that maximizes the payoff of all players is not $\overline{N}$, but the singleton structure: $\mathcal{N}^* = \underline{N}$.

Using $\mathcal{N}^* = \underline{N}$, we find that the coalitional game generated by the projection rule is given by $w_{pr}^{\underline{N}}(S) = 4$ if $s = 1$, $w_{pr}^{\underline{N}}(S) = 9$ if $s = 2$, $w_{pr}^{\underline{N}}(S) = 12$ if $s = 3$, and $w_{pr}^{\underline{N}}(N) = 16$.

Because the partition function game v has negative externalities, the pessimistic rule assigns $p(S, \mathcal{S}, \mathcal{N}, v) = \overline{S^c}$, so that $w_p(S) = 0$ if $s = 1$, $w_p(S) = 3$ if $s = 2$, $w_p(S) = 12$ if $s = 3$, and $w_p(N) = 16$.

A necessary and sufficient condition for non-emptiness of the core of a symmetric coalitional game w is that the game is top-convex (Jackson and van den Nouweland (2005)), i.e., $\frac{w(N)}{n} \geq \frac{w(S)}{s}$ for all $S \subseteq N$. Notice that this condition is satisfied for the pessimistic game w_p but not for the projection game $w_{pr}^{\underline{N}}$. Hence, the projection core is empty, whereas the pessimistic core is not.

The optimistic core is a subset of the projection core and is thus also empty for the partition function game in this example. We point out that the optimistic game w_o coincides with the projection game in this example, which follows from the negative externalities (which imply that the optimistic expectations are $o(S, \mathcal{S}, \mathcal{N}, v) = \underline{S^c}$).

6 Applications

In this section, we compute the projection core and the optimistic and pessimistic cores in two applications of coalitional games with externalities, which have been emphasized in Industrial Organization and Environmental Economics. Because we single out the projection core as the best stability concept for games with externalities, our objective is primarily to provide conditions under which the projection core is nonempty. We focus on settings where players are ex ante symmetric, and assume specific functional forms to simplify the expressions of partition functions.

6.1 Cartels

The first application considers the formation of cartels (or mergers by firms) in a linear Cournot market. We consider the version of the model studied by Bloch (1996) and Ray and Vohra (1999). Consider n firms in an oligopoly with linear inverse demand

$$P = 1 - \sum_{i=1}^{n} q_i.$$

Variable costs are normalized to 0, but each coalition of firms incurs a fixed cost K of operating on the market. Suppose that firms merge so that the final market structure is given by the coalition structure $\mathcal{N} = \{S_1, ..., S_M\}$. The firms in each coalition S_m select one representative, which will be the active firm on the market. Computations of the equilibrium of the Cournot quantity setting game easily show that the profit of coalition S_m only depends on the number M of coalitions in the market and is given by

$$v(S_m, \mathcal{N}) = \max\{\frac{1}{(M+1)^2} - K, 0\},$$

yielding a partition function game with *positive externalities*. In order to compute the optimal partition $\mathcal{N}^*$, notice that the sum of profits net of fixed costs in a market with M coalitions is given by $\frac{M}{(M+1)^2}$ whereas the profit of a monopolist is $\frac{1}{4}$. It is easy to check that $\frac{M}{(M+1)^2} < \frac{1}{4}$ for any $M \geq 2$, so that the coalition structure that maximizes the sum of profits of all firms is $\mathcal{N}^* = \overline{N}$. This fact, together with positive externalities, implies that the projection rule *pr* is equivalent to the optimistic rule and $pr(S, \mathcal{S}, \overline{N}, v) = \overline{S^c}$ for all S and $\mathcal{S}$. The group of firms S will not necessarily form a single coalition, but will choose the optimal number t^* of coalitions to maximize the sum of profits of its members. Thus, we compute the coalitional function generated by the projection rule as[7]

$$w_{pr}(S) = \max_{1 \leq t \leq s} \frac{t}{(t+2)^2} - tK.$$

Let $g(t) = \frac{t}{(t+2)^2} - tK$. As $\frac{\partial g}{\partial t} = \frac{2-t}{(t+2)^3} - K$, the optimal number of coalitions in S is either $t^* = 1$ or $t^* = 2$. For any S with $2 \leq s \leq n-1$, we find $t^* = 2$ if $K \leq \frac{1}{72}$ and $t^* = 1$ if $K \geq \frac{1}{72}$. Notice that when $K > \frac{1}{9}$, the maximum obtainable profit of group S is negative, so that the members of S are better

[7]Throughout the remainder of this section we write w_{pr} instead of $w_{pr}^{\overline{N}}$.

off not producing. Summarizing, we obtain:

$$\text{If } K \leq \frac{1}{72} \qquad \begin{aligned} w_{pr}(S) &= \frac{1}{9} - K \text{ if } s = 1, \\ w_{pr}(S) &= \frac{1}{8} - 2K \text{ if } 2 \leq s \leq n-1, \\ w_{pr}(N) &= \frac{1}{4} - K. \end{aligned}$$

$$\text{If } \frac{1}{9} \geq K \geq \frac{1}{72} \qquad \begin{aligned} w_{pr}(S) &= \frac{1}{9} - K \text{ if } 1 \leq s \leq n-1, \\ w_{pr}(N) &= \frac{1}{4} - K. \end{aligned}$$

$$\text{If } K \geq \frac{1}{9} \qquad \begin{aligned} w_{pr}(S) &= 0 \text{ if } 1 \leq s \leq n-1, \\ w_{pr}(N) &= \frac{1}{4} - K. \end{aligned}$$

The core of this symmetric coalitional game is nonempty if and only if the game is top-convex (cf. Jackson and van den Nouweland (2005)), i.e.

$$\frac{w_{pr}(N)}{n} \geq \frac{w_{pr}(S)}{s} \text{ for all } S \subset N. \tag{11}$$

Because

$$\max_{1 \leq s \leq n-1} \frac{w_{pr}(S)}{s} = \frac{1}{9} - K,$$

condition (11) becomes:

$$K \geq \frac{4n - 9}{36(n-1)}.$$

Hence, the projection core of the game is nonempty if and only if the fixed cost is larger than an upper bound that tends to $\frac{1}{9}$ when n goes to infinity. The same holds for the optimistic core, which coincides with the projection core.

Consider now the pessimistic core. Given that the game has positive externalities, the pessimistic rule is $p(S, \mathcal{S}, \overline{N}, v) = \underline{S^c}$ for all S and $\mathcal{S}$. If the group S forms t coalitions, the total number of coalitions in the market is $t+n-s$. The coalitional game generated by the pessimistic rule is thus given by

$$w_p(S) = \max_{1 \le t \le s} \frac{t}{(t+n-s+1)^2} - tK.$$

Let $t^*(s,K)$ be the optimal number of coalitions chosen by group S when the fixed cost is K and assume that K is low enough so that members of S produce a positive quantity.[8] Then for any $S \subset N$

$$\begin{aligned} \frac{w_p(S)}{s} &= \frac{t^*(s,K)}{s(t^*(s,K)+n-s+1)^2} - \frac{t^*(s,K)K}{s} \\ &\le \frac{t^*(s,K)}{s(t^*(s,K)+n-s+1)^2} - \frac{K}{s} \\ &< \frac{t^*(s,K)}{s(t^*(s,K)+n-s+1)^2} - \frac{K}{n}. \end{aligned} \tag{12}$$

where the first inequality stems from the fact that $t^*(s,K) \ge 1$ and the second from the fact that $s < n$. Next define $h(t) \equiv \frac{t}{(t+n-s+1)^2}$. It is easy to check that

$$\frac{\partial h(t)}{\partial t} = \frac{n-s+1-t}{(t+n-s+1)^3}.$$

Thus, for a fixed s, $\frac{\partial h(t)}{\partial t}$ is positive when $t < n-s+1$ and negative when $t > n-s+1$, so that $h(t)$ attains its maximum at $\hat{t}(s) = n-s+1$. It follows that

$$\begin{aligned} \frac{t^*(s,K)}{s(t^*(s,K)+n-s+1)^2} &= \frac{h(t^*(s,K))}{s} \\ &\le \frac{h(\hat{t}(s))}{s} \\ &= \frac{n-s+1}{4s(n-s+1)^2} = \frac{1}{4s(n-s+1)}. \end{aligned} \tag{13}$$

Furthermore, $s(n-s+1)-n = (n-s)(s-1) \ge 0$ for any $1 \le s < n$, so that

[8] If the members of S do not produce, then top convexity holds trivially.

$$\begin{aligned}\frac{w_p(N)}{n} &= \frac{1}{4n} - \frac{K}{n} \geq \frac{1}{4s(n-s+1)} - \frac{K}{n} \\ &\geq \frac{t^*(s,K)}{s(t^*(s,K)+n-s+1)^2} - \frac{K}{n} > \frac{w_p(S)}{s},\end{aligned}$$

where the second inequality stems from (13) and the third from (12). The game w_p is thus top convex, establishing that the pessimistic core is nonempty for any value of the fixed cost K.

6.2 Public goods

In the second application, we consider a model of public good provision inspired by Ray and Vohra (2001)'s study of global environmental public goods. Different from Ray and Vohra (2001), we do not assume that the public good is necessarily global, and we introduce a parameter measuring the degree of spillovers between local public goods provided in different jurisdictions. (This formulation was introduced by Bloch and Zenginobuz (2006).) Let $\mathcal{N} = \{S_1, ..., S_M\}$ be the coalition structure representing the different jurisdictions (or countries). The utility function of agent i in jurisdiction m is given by:

$$U_i = g_m + \alpha \sum_{j \neq m} g_j - \frac{1}{2} c_i^2,$$

where g_j is the amount of public good provided in jurisdiction j, $\alpha \in [0,1]$ a spillover parameter, and c_i the contribution of agent i. Notice that when $\alpha = 1$ the public good is a pure (or global) public good whereas when $\alpha = 0$, the public good is a local public good. We suppose that the contribution level of agents in jurisdiction j is jointly chosen in order to maximize the sum of utilities of all agents in the jurisdiction. Given the convexity of the contribution cost function, it is optimal for all agents in a jurisdiction j to contribute the same amount c_j and thus $g_j = s_j c_j$. The optimal contributions level in jurisdiction j is then $c_j^* = s_j$, so that the total public good contribution in the jurisdiction is $g_j^* = s_j^2$. Substituting in the sum of utilities of agents in each jurisdiction, we derive the partition function

$$v(S_m, \mathcal{N}) = s_m(\frac{s_m^2}{2} + \alpha \sum_{j \neq m} s_j^2).$$

For any two disjoint jurisdictions S_l and S_m in $\mathcal{N}$, we have

$$\begin{aligned}
&v(S_l \cup S_m, \mathcal{N} \setminus \{S_l, S_m\} \cup \{S_l \cup S_m)\}) \\
&\qquad = \frac{(s_l + s_m)^3}{2} + \alpha(s_l + s_m) \sum_{j \neq l,m} s_j^2 \\
&\qquad = \frac{s_l^3}{2} + \frac{s_m^3}{2} + \frac{3}{2} s_l^2 s_m + \frac{3}{2} s_l s_m^2 + \alpha(s_l + s_m) \sum_{j \neq l,m} s_j^2 \\
&\qquad > \frac{s_l^3}{2} + \frac{s_m^3}{2} + \alpha s_l^2 s_m + \alpha s_l s_m^2 + \alpha(s_l + s_m) \sum_{j \neq l,m} s_j^2 \\
&\qquad = \frac{s_l^3}{2} + \alpha s_l \sum_{j \neq l} s_j^2 + \frac{s_m^3}{2} + \alpha s_m \sum_{j \neq m} s_j^2 \\
&\qquad = v(S_l, \mathcal{N}) + v(S_m, \mathcal{N}),
\end{aligned}$$

showing that the partition function game is superadditive. In addition, for any $S_k, S_l, S_m \in \mathcal{N}$,

$$\begin{aligned}
v(S_m, \mathcal{N} \setminus \{S_k, S_l\} \cup \{S_k \cup S_l)\}) &= \frac{s_m^3}{2} + \alpha s_m (s_k + s_l)^2 + \alpha s_m \sum_{j \neq k,l,m} s_j^2 \\
&> \frac{s_m^3}{2} + \alpha s_m s_k^2 + \alpha s_m s_l^2 + \alpha s_m \sum_{j \neq k,l,m} s_j^2 \\
&= v(S_m, \mathcal{N}),
\end{aligned}$$

so that the game has positive externalities. As we saw in Proposition 5.4, for this superadditive game with positive externalities, $\mathcal{N}^* = \overline{N}$, the coalitional function $w_{pr}^{\overline{N}}$ generated by the projection rule is equal to the coalitional function w_o generated by the optimistic rule, and

$$w_{pr}^{\overline{N}}(S) = v(S, \overline{S} \cup \overline{S^c}) = \frac{s^3}{2} + \alpha s (n-s)^2.$$

Because this game is symmetric, a necessary and sufficient condition for its core to be nonempty is that it is top convex, i.e.,

$$\frac{w_{pr}^{\overline{N}}(N)}{n} = \frac{n^2}{2} \geq \frac{w_{pr}^{\overline{N}}(S)}{s} = \frac{s^2}{2} + \alpha(n-s)^2 \text{ for all } S \subset N.$$

We verify that the function $g(s) = \frac{s^2}{2} + \alpha(n-s)^2$ is strictly convex so that it achieves a maximum at $s = n$ if and only if

$$g(1) = \frac{1}{2} + \alpha(n-1)^2 \leq \frac{n^2}{2} = g(n).$$

Hence, the projection core of the public good game is nonempty if and only if $\alpha \leq \frac{n+1}{2(n-1)}$. We thus observe that the projection core of the public good game is nonempty if and only if the level of spillovers across jurisdictions is bounded above by a parameter that converges to $\frac{1}{2}$ when n goes to infinity. The same holds for the optimistic core, which coincides with the projection core.

Given that the game has positive externalities, the pessimistic rule is $p(S, \mathcal{S}, \overline{N}, v) = \underline{S^c}$ for all S and $\mathcal{S}$, and the coalitional function generated by the pessimistic rule is:

$$w_p(S) = \frac{s^3}{2} + \alpha s(n-s).$$

Define $h(s) = \frac{w_p(S)}{s} = \frac{s^2}{2} + \alpha(n-s)$. The pessimistic core is nonempty if and only if the game w_p is top convex, which is equivalent to $h(s)$ attaining a maximum at n. Because the function $h(s)$ is strictly convex, it $h(s)$ attains a maximum at n if and only if

$$h(1) = \frac{1}{2} + \alpha(n-1) \leq \frac{n^2}{2} = h(n),$$

a condition that is satisfied for all $\alpha \leq 1$ and $n \geq 1$. Hence, the pessimistic core is nonempty for all values of α.

7 Conclusion

This paper proposes axiomatic foundations to expectation formation rules, by which deviating players anticipate the reaction of external players in a partition function game. We single out the projection rule – where players anticipate that external players project the current partition – as the only rule satisfying subset consistency and responsiveness or subset consistency, independence of the original partition of deviating players, and coherence of expectations. Exogenous rules are the only rules satisfying subset consistency and independence of the original partition, and the pessimistic rule is the only rule among the common rules proposed in the literature that satisfies preservation of superadditivity. Our axiomatic analysis suggests that the projection core is a natural candidate to study the stability of games in partition function form, and we analyze the projection core in two standard applications of coalition formation with externalities.

One of the major drawbacks of our analysis (as of any analysis of the core) is that we only consider myopic deviations, and do not describe the full process by which coalitions successively deviate from an allocation. In particular, we do not submit the allocation of deviating players to the same stability test as the original allocation. In the context of partition function games, the recursive core studied by Huang and Sjostrom (2003) and Koczy (2007) captures this requirement, by assuming that deviating players anticipate that external players will select a point in the core of the game restricted to external players. Alternatively, one could consider farsighted players, as in Chwe (1994) or Diamantoudi and Xue (2003), and analyze the farsighted core of the game generated by different expectation formation rules. When expectation formation rules are independent of the original partition, we suspect that the analysis of the farsighted core is a straightforward extension of the myopic core. When expectation formation rules are responsive to the original partition, as in the case of the projection rule, the analysis of the farsighted core involves a dynamic process of expectation formation, and we hope to undertake such an analysis in future research.

8 References

Aumann, R. (1967) "A survey of cooperative games without side payments," in *Essays in Mathematical Economics*, M. Shubik ed., Princeton: Princeton University Press.

Bloch, F. (1996) "Sequential formation of coalitions in games with externalities and fixed payoff division," *Games and Economic Behavior* **14**, 90-123.

Bloch, F. and U. Zenginobuz (2006) "Tiebout equilibria in local public good economies with spillovers," *Journal of Public Economics* **90**, 1745-1763.

Chander, P. and H. Tulkens (1997) "The core of an economy with multilateral environmental externalities," *International Journal of Game Theory* **26**, 379-401.

Chwe, M. (1994) "Farsighted coalitional stability," *Journal of Economic Theory* **63**, 299-325.

De Clippel, G. and R. Serrano (2008) "Marginal contributions and externalities in the value", *Econometrica* **76**, 1413-1436.

Diamantoudi, E. and L. Xue (2003) "Farsighted stability in hedonic games," *Social Choice and Welfare* **21**, 39-61.

Dutta, B., Ehlers, L. , Kar, A. (2010) "Externalities, potential, value and consistency, " *Journal of Economic Theory* **145**, 2380-2411.

Hart, S. and M. Kurz (1983) "Endogenous formation of coalitions," *Econometrica* **51**, 1047-1064.

Huang, C.Y. and T. Sjostrom (2003) "Consistent solutions for cooperative games with externalities," *Games and Economic Behavior* **43**, 193-216.

Jackson, M. and A. van den Nouweland (2005) "Strongly stable networks," *Games and Economic Behavior* **51**, 420-444.

Koczy, L. (2007) "A recursive core for partition function form games," *Theory and Decision* **63**, 41-51.

Macho-Stadler, I., D. Perez-Castrillo, D. Wettstein (2007) "Sharing the surplus: An extension of the Shapley Value for environments with externalities," *Journal of Economic Theory* **135**, 339-356.

Myerson, R. (1977) "Values of games in partition function form", *International Journal of Game Theory* **6**, 23-31.

Ray, D. (2007) *A Game Theoretic Perspective on Coalition Formation*, Oxford: Oxford University Press.

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.

Shenoy, P. (1979) "On coalition formation: A game theoretical approach," *International Journal of Game Theory* **8**, 133-164.

Thrall, R.M. and W.F. Lucas (1963) "n-Person games in partition function form", *Naval Research Logistics Quarterly* **10**, 281-298.

Von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.

A Verifications of properties of common expectation formation rules

We first consider the disintegration rule. This rule satisfies path independence because for any deviating coalition S it predicts that all blocks U in $\mathcal{N}$ that do not intersect S remain intact, while the other blocks disintegrate into singletons. This is obviously independent of the order in which the members of S form expectations. The disintegration rule also satisfies coherence of expectations because in the partition $\mathcal{S} \cup f(S, \mathcal{S}, \mathcal{N}, v)$ the blocks in $\mathcal{S}$ do not intersect the blocks in $f(S, \mathcal{S}, \mathcal{N}, v)$. The disintegration rule violates all other axioms:

Example A.1 *(The disintegration rule.)*

Let $N = \{i, j, k, l\}$ and let f be the disintegration rule. Because this rule does not depend on $\mathcal{S}$ or v, we simplify notation and write $f(S, \mathcal{N})$.

IPDOP (and thus IOP) is violated because for $\mathcal{N} = \{ijkl\}$ and $\mathcal{N}' = \{ij|kl\}$, and $S = \{i, j\}$, we have that $\mathcal{N}|_{S^c} = \mathcal{N}'|_{S^c}$, while $f(S, \mathcal{N}) = \{k|l\} \neq \{kl\} = f(S, \mathcal{N}')$.

RPEOP is violated because for $\mathcal{N} = \{ijkl\}$ and $\mathcal{N}' = \{i|j|k|l\}$, and $S = \{i, j\}$, we have that $\mathcal{N}|_{S^c} \neq \mathcal{N}'|_{S^c}$, while $f(S, \mathcal{N}) = \{k|l\} = f(S, \mathcal{N}')$.

Subset consistency is violated because for $S = \{i, j\}$, $T = \{i\}$, and $\mathcal{N} = \{i|jkl\}$, we have that $f(S, \mathcal{N}) = \{k|l\}$, while $f(T, \mathcal{N}) = \{jkl\}$.

PSA is violated for the partition function game in Example 4.8 when $\mathcal{N} = \{i|jkl\}$ because $w_f^{\mathcal{N}}(i) + w_f^{\mathcal{N}}(jk) = v(i, i|jkl) + v(jk, i|jk|l) = 4 + 7 > 8 = v(ijk, ijk|l) = w_f^{\mathcal{N}}(ijk)$.

We have already established that the projection rule satisfies IPDOP, RPEOP, PI, SC, and COH. The rule violates PSA for the game in example 4.8 when $\mathcal{N} = \{i|jkl\}$, because $w_f^{\mathcal{N}}(i) + w_f^{\mathcal{N}}(jk) = v(i, i|jkl) + v(jk, i|jk|l) = 4 + 7 > 8 = v(ijk, ijk|l) = w_f^{\mathcal{N}}(ijk)$.

The $\mathcal{M}$-exogenous projection rules satisfy IOP, and thus PI. Because these rules satisfy SC and IPDOP, by Theorem 4.4 they violate COH.

The pessimistic, optimistic, and max rules satisfy IOP and thus also PI. By theorem 4.6, they thus violate SC. The following example demonstrates that these rules violate COH.

Example A.2 *(Violations of coherence.)*

Let $N = \{1, 2, 3, 4\}$ *and consider a partition function game in which* $v(1|2|34) = (4, 4, 4)$, $v(12|34) = (6, 10)$, $v(1|2|3|4) = (1, 1, 1, 1)$, $v(12|3|4) = (5, 3, 3)$.

Then, for the optimistic rule o, *and* $S = \{12\}$, $\mathcal{S} = \{1|2\}$, *we have* $o(S, \mathcal{S}, v) = \{34\}$ *and* $o(S^c, o(S, \mathcal{S}, v), v) = o(34, \{34\}, v) = \{12\} \neq \mathcal{S}$, *demonstrating that the optimistic rule violates COH.*

For the pessimistic rule p, *and* $S = \{12\}$, $\mathcal{S} = \{12\}$, *we have* $p(S, \mathcal{S}, v) = \{3|4\}$ *and* $p(S^c, p(S, \mathcal{S}, v), v) = p(34, \{3|4\}, v) = \{1|2\} \neq \mathcal{S}$, *demonstrating that the pessimistic rule violates COH.*

The max rule m *violates COH because when* $S = \{12\}$, $\mathcal{S} = \{12\}$, *we have* $m(S, \mathcal{S}, v) = \{34\}$ *and* $m(S^c, m(S, \mathcal{S}, v), v) = m(34, \{34\}, v) = \{1|2\} \neq \mathcal{S}$.