Preliminary and Incomplete

Sequential Bargaining, Coalitions, and Externalities*

E. Maskin[a], G. de Clippel[b], and R. Serrano[b]

April 2021

Abstract

We argue that cooperative game theory has failed to realize its early promise for positive economics because its most prominent solution concepts, the core (Gillies 1959) and the Shapley value (Shapley 1953), unrealistically predict that the grand coalition – the coalition of all players – will always form. Not only is this prediction inaccurate empirically, it is, we contend, dubious theoretically. Specifically, we propose that when coalitions arise though sequential bargaining and there are significant externalities between coalitions, the grand coalition should typically *not* be predicted to form because smaller coalitions will benefit from free riding. One way to interpret this logic is as a refutation of the Coase Theorem (1960) when there are more than two agents.

We develop a generalization of the Shapley value that formalizes our approach. Our generalized Shapley value is defined axiomatically, but the axioms are quite different from Shapley's own axiomatization (in fact, they are similar in form to those of the core). We also provide several non-cooperative implementations of our solution concept, thereby carrying out Nash's program (Nash 1953). As an application of the theory, we offer a new explanation for Zipf's Law for firms.

* This is a much revised version of a paper prepared by the first author for his presidential address to the Econometric Society (Maskin 2003).

[a] Harvard University and Higher School of Economics
[b] Brown University

1. Introduction

For the last forty years, cooperative game theory (the theory of how coalitions – groups of agents that act in concert – behave) has been dominated by its non-cooperative counterpart, at least judging from their respective impacts on mainstream positive economics.[1] That cooperative theory should be overshadowed seems regrettable because its defining approach – making *coalitions* central – offers potential insight into many aspects of economic and political life, from the European Union to the Paris Accord on Climate Change, to the OPEC cartel. Indeed, if we conceive of an economic institution, e.g., a firm, as a coalition of its members, then cooperative game theory holds out the possibility of explaining why we see the institutions we do.

We believe that one reason that cooperative theory has not had more influence on economics is that its two most visible solution concepts for games of three or more players, the core (Gillies 1959) and the Shapley value (Shapley 1953), normally *imply* that the grand coalition – the coalition of all players – forms.[2] This is clearly empirically unrealistic, e.g., the European Union is not the only political coalition in the world. Moreover, by positing the grand coalition, the analyst rules out, from the beginning, the possibility of interaction *between* institutions – the focus of most economics.

Indeed, we maintain, there are many settings in which theory should predict that the grand coalition will *not* form. For example, imagine that when its members expend effort, a coalition can produce a public good – a good whose benefits cannot be confined to the coalition alone. In these circumstances, a player outside the coalition may do better by remaining there –

[1] Three-quarters of von Neumann and Morgenstern's (1944) pioneering treatise on game theory was devoted to cooperative games. By contrast, Fudenberg and Tirole (1991), a leading modern game theory textbook, contains only non-cooperative theory, i.e., no cooperative games at all.
[2] The core predicts the grand coalition when it makes a prediction at all. But in many games, the core is empty, as we illustrate in Example 1.

and enjoying the fruits of the public good without exerting effort – rather than by joining it. Of course, this conclusion depends on the process by which coalitions form. We will argue that one should typically expect multiple coalitions when they arise through *sequential bargaining.* That is, we imagine that rather than all agents sitting down simultaneously to hash out an agreement, all negotiations are bilateral, with additional agents coming in one at a time (the order in which this occurs is assumed to be random). We believe this sequentiality reflects an important feature of reality: that large institutions are usually formed over time through a succession of bilateral bargains.[3]

Coase (1960) reasoned that, even in the presence of externalities like public goods, unconstrained bargaining should result in an efficient outcome. Coase himself gave an argument for this conclusion only in the case of two agents (and it was an informal argument at that). But others have applied the Coase Theorem to unlimited numbers. Our results suggest that the limitation to two was perhaps justified after all. (Another possible interpretation is that requiring bargaining to be sequential is itself a constraint on bargaining.[4])

To model a sequential bargaining process, we propose a generalization of the Shapley value as our solution concept, i.e., as a prediction for which coalitions will form and what players' payoffs will be.

When there are no externalities between coalitions (formally modeled by assuming the game can be described by a characteristic function), we predict that players receive expected payoffs given by the ordinary Shapley value (Theorem 1). Our axioms characterizing this solution, however, are quite different from Shapley's own (interestingly, they are similar in form

---

[3] For example, the corporate giant Facebook has grown in large part by successively buying up other companies, e.g., Instagram and WhatsApp.

[4] One problem with understanding the Coase Theorem is that neither Coase nor most of his followers make the premises precise.

to those defining the core[5]). Instead, they are formulated to capture the essence of sequential bargaining.[6]

When there are externalities between coalitions (and the game is formally described by a partition function), our axioms must be generalized; in particular, they must accommodate the possibility of a non-degenerate coalition structure (an outcome other than the grand coalition). Even so, the prediction they lead to is generically unique (Theorem 2). That is, except in knife-edge cases, there is only one coalition structure and only one payoff vector consistent with the axioms. Indeed, if it turns out that there are no externalities, the generalized axioms predict that the grand coalition will form and that payoffs are given by the ordinary Shapley value (Theorem 3).

When externalities between coalitions are negative, then an outside player may have an additional motivation to join a coalition: avoiding negative payoffs. Indeed, when there are only three players, we establish that given our axioms, the grand coalition always forms with negative externalities (Theorem 4). Interestingly, this result does not extend to games of four or more players, as we show by example.

Nash (1953) argued that game theory should consist of an interplay between cooperative and non-cooperative analysis. For each cooperative game – for which details such as players' strategies and timing are typically not specified explicitly – there will be a variety of non-cooperative "implementations" in which the omitted specifications are filled in. Thus, we get the benefit of both the "general" and the "specific." This dual approach is called the Nash Program, and we will follow it here by exhibiting several explicit non-cooperative games whose subgame

[5] They also resemble Hart and Mas-Colell (200_)'s axiomatization of the Shapley value.
[6] Despite, its axiomatic similarity to the Shapley value, the core, does not model sequential bargaining well, as we argue in Section 2.

perfect equilibrium outcomes coincide with those of the generalized Shapley value (Theorems 5 and 6).

To illustrate how our generalized Shapley value can be used for applications, we re-examine Zipf's Law, a celebrated empirical regularity from industrial organization. The Law asserts that the distribution of firm sizes – in a country or an industry – is given by a power law: the number of firms bigger than or equal to *s* is inversely proportional to *s*.

Zipf's Law is usually explained as a consequence of firm growth: it holds in the long run if firms' growth rates are drawn independently from the same distribution (see Gabaix 1999). Our approach offers an alternative story. Suppose that the merger of two firms gives them higher profit than before but also exerts an even bigger positive externality on other firms.[7], We show that, under such circumstances, the generalized Shapley value predicts a power-law distribution of firm sizes (Theorems 7 and 8), at least in simple cases.

We begin, in Section 2, by studying two simple examples that illustrate our approach. In Section 3, we briefly treat the model with no externalities, and develop the general theory allowing for externalities in Section 4. We discuss non-cooperative implementation in Section 5, and Zipf's Law in Section 6. Finally, in Section 7, we draw connections to the literature.

2. Two Examples

A solution concept for an *n*-player cooperative game is a prediction of players' payoffs (usually expressed as a set of payoff *n*-tuples) and of the coalitions they form (most existing solution concepts predict the grand coalition). One leading solution concept is the *core*, defined as the set of players' payoff *n*-tuples that (i) are *feasible* (i.e., the sum of payoffs is no greater

[7] This externality may come about because a bigger firm typically has greater market power, which allows it to set higher prices – and thereby gives its competitors this opportunity too.

than the maximum possible payoff for the grand coalition[8]) and (ii) cannot be *blocked* by any coalition (a coalition C blocks a payoff n-tuple if the sum of the payoffs to the players in C is less than the maximum payoff that C can get on its own).[9]

Unfortunately, the core fails to make a prediction at all for many games, i.e., it is often empty. Here is such an example:

Example 1: A Game of Production

Consider a game of three players, 1, 2, and 3. The coalition of players 1 and 2 can produce 16; players 1 and 3 together can produce 17; and players 2 and 3 together can produce 18. Finally, the grand coalition can produce 25, and each player on her own can produce 6.[10]

The first thing notice about this game is that its core is empty. If (x_1, x_2, x_3) is a triple of payoffs, then, to be in the core, it must satisfy

(1) $$x_1 + x_2 \geq 16$$

otherwise the coalition $\{1,2\}$ can block[11]; similarly,

(2) $$x_1 + x_3 \geq 17$$

and

(3) $$x_2 + x_3 \geq 18$$

[8] Throughout the paper we concentrate on games with *transferable utility*. Implicitly, there exists some divisible and tradeable good (e.g., money) in plentiful supply, and a player receiving transfer Δ of this good (where Δ can be positive or negative) experiences a payoff change of Δ (i.e., his payoffs is linear in money). Hence, we can speak of a *coalition's payoff*: it is simply the sum of the payoffs for the players in the coalition.
[9] See Section 3 for a formal definition/ axiomatization of the core.
[10] The units for these numbers are dollars -- or whatever else the transferable good may be.
[11] By "blocking" we mean that {1,2} can, in effect, go off by themselves and attain payoffs that sum to 16.

to prevent coalitions $\{1,3\}$ and $\{2,3\}$, respectively, from blocking. But adding up (1), (2), and (3), we obtain $2(x_1+x_2+x_3) \geq 51$, i.e.,

$$x_1+x_2+x_3 \geq 51/2,$$

which violates the fact that the grand coalition can produce at most $25(=50/2)$.

Despite the emptiness of the core, we will argue that a clear-cut outcome arises in this game if it derives from sequential bilateral bargaining and binding agreements.

Specifically, suppose that player 1 arrives at the "bargaining site" first, followed by 2 and finally 3. Because 1 and 2 get there ahead of 3, we imagine that they will first make a binding agreement between themselves and that when 3 arrives, the coalition $\{1,2\}$ will bargain with her.

What will happen when 3 shows up? If she remains on her own, she gets a payoff of 6. Thus, $\{1,2\}$ must offer her (slightly more than) 6 to induce her to join them (and doing so is in their interest because, if she joins them, she contributes $25-16=9$, which exceeds 6). We conclude that

(4) $$\varphi_3 = 6,$$ [12]

where φ_i [13] denotes the predicted payoff for player *i*, her payoff according to our solution concept.

Moving backward in time, let us consider what player 2 must offered to get her to sign an agreement with 1. If 2 fails to be lured, then presumably she will sign an agreement with 3 (otherwise, she is stuck with a payoff of 6), in which case her payoff φ_2 satisfies

[12] In effect, we give all bargaining power to coalition {1,2} (and, more generally, to earlier-arriving coalitions). But this asymmetry is eliminated when we randomize over the order of arrival below.

[13] We use the notation "φ_i" – the traditional symbol for the Shapley value – because of the connection we will draw between that solution concept and our own.

(5) $\varphi_2 + \varphi_3 = 18$,

where 18 is the total payoff that $\{2,3\}$ can attain. From (4), $\varphi_3 = 6$ [14] and so $\varphi_2 = 12$, which is thus also what player 2 must receive to join 1.[15] To derive player 1's payoff, we have, from feasibility,

(6) $\varphi_1 + \varphi_2 + \varphi_3 = 25$

Equations (4) - (6) determine players' payoffs:

(7) $(\varphi_1, \varphi_2, \varphi_3) = (7, 12, 6)$

Note that player 1 is willing to do business with 2 because the payoff she gets as a result – 7 – is bigger than what she would get alone, i.e., 6. That is, the prediction that the grand coalition forms make sense.

The order of arrival here – 1, 2, and then 3 – is purely accidental. Any of the other five orders are equally possible. But if we perform the corresponding calculations in those other cases, we continue to expect the grand coalition to form. Moreover, if we average over (7) and its five counterparts, the three players receive exactly their *Shapley values*, $\left(\frac{47}{6}, \frac{50}{6}, \frac{53}{6}\right)$[16].

Our analysis of Example 1 illustrates that the Shapley value epitomizes outcomes that result from sequential bargaining and binding bilateral agreements. By contrast, the core is at odds with such outcomes. In our sequential scenario, by the time player 3 enters the negotiations, 2 is already bound to 1 by contract. And so the fact that 2 and 3 hypothetically could jointly

[14] We derived $\varphi_3 = 6$ in the case that $\{1,2\}$ bargains with 3, but it also holds when 2 bargains with 3.

[15] Note that 12 exceeds payoff 6, which is the best player 2 could get on her own. Thus, she will be willing to join either player 3 or 1.

[16] To see why the Shapley values obtain if we average over players' payoffs from the six arrival orders, note that (7) gives players their marginal contributions: player 3 adds payoff 6 by joining the null coalition; player 2 adds 12 by joining $\{3\}$, and player 1 adds 7 in joining $\{2,3\}$. And, as Roth (1977) demonstrates, players' Shapley values are just their expected marginal contributions, where the expectation is taken over all possible orders in which the grand coalition can be assembled.

produce 18 – which is pertinent for the core – is no longer relevant to bargaining. In particular, it has no bearing on the compensation 3 received for joining $\{1,2\}$ (by contrast, it *does* bear on 2's compensation in her dealings with 1).

The game in Example 1 has the property that a coalition's maximum payoff is independent of which other coalitions form (for example, player 3 can get 6 by herself, regardless of whether 1 and 2 get together); this property is what we mean by "no externalities." Theorem 3 in Section 4 establishes that as long as the property holds, the grand coalition can be predicted to form. However, once that property is relaxed, such a prediction may no longer be justified, as the next example illustrates.

Example 2: A Public Good Game[17]

Suppose that, working in different combinations, players 1, 2, and 3 can produce a public good, whose level depends on which coalition has formed. Specifically, the coalition $\{1,2\}$ can produce 12; $\{1,3\}$ can produce 13, and $\{2,3\}$ can produce 14. The grand coalition can produce 24. A player produces nothing on his own, but if the other two players form a coalition, he can free-ride on the public good they produce and enjoy a payoff of 9.[18]

We claim that one should not expect the grand coalition to form in this game if the outcome is reached through sequential bargaining. To see why not, consider the player arrival order 1, 2, and then 3. If players 1 and 2 have formed a coalition before 3 arrives, then they must offer him 9 to induce him to join too, since he can derive this payoff by free-riding on them:

(8) $$\varphi_3 = 9.$$

[17] This game is closely related to Example 1.2 in Ray and Vohra (1999).

[18] These numbers collectively specify the partition-function form of the game, a generalization of the characteristic function form that allows for externalities between coalitions. For more details see Section 4.

Now, earlier in time, let's consider what player 2 must be offered to join 1. If he *doesn't* join 1, he could presumably try to sign up 3. Notice, in fact, that he would need to offer player 3 only epsilon above 0, since in this case, there would be no coalition $\{1,2\}$ for 3 to free-ride on. Hence, from forming coalition $\{2,3\}$, player 2 gets φ_2, where

(9) $$\varphi_2 + \varphi_3' = 14$$

and

(10) $$\varphi_3' = 0,$$

where the right-hand side of (9) reflects $\{2,3\}$'s ability to produce 14, and the right-hand side of (10) embodies player 3's inability to free-ride. That is,

(11) $$\varphi_2 = 14$$

(and so joining 3 is better for player 2 than remaining alone and getting payoff 0). Thus, player 2 must get 14 to be induced to join 1.

From (8)-(11), player 1's net payoff if 2 and 3 are both enlisted in a coalition is

(12) $$24 - 14 - 9 = 1,$$

and his net payoff if only 2 joins him is

(13) $$12 - 14 = -2.$$

Alternatively, he can refrain from signing up either player. In that case, as we saw above, the coalition $\{2,3\}$ can be expected to form, and player 1 enjoys a free-riding payoff of

(14) $$\varphi_1 = 9,$$

higher than the payoffs from his other options.

Thus, with arrival order 1, 2, 3, we predict that *two* coalitions will form: {2, 3} and {1}. The resulting payoffs are, from (8) - (11) and (14),

(15) $(\varphi_1, \varphi_2, \varphi_3') = (9, 14, 0).$

The failure of the grand coalition to come about here is clearly due to the externalities. Specifically, player 3 requires a big payoff to join coalition $\{1, 2\}$ because he can otherwise free-ride on them. But this payoff – together with the hefty payoff that player 2 commands – leaves so little for player 1 that he is better off free-riding on $\{2, 3\}$.

For each other possible arrival order, we obtain an analogous outcome consisting of two coalitions. More precisely, we predict that, with 1/3 probability, we get each of the following coalition structures:

$$\{\{1\},\{2,3\}\}, \ \{\{2\},\{1,3\}\}, \ \text{and} \ \{\{3\},\{1,2\}\}.$$

As in Example 1, we can calculate the Shapley value (more accurately, the "generalized Shapley value") for each player by averaging over these orders. We thereby obtain $\left(\frac{43}{6}, \frac{44}{6}, \frac{45}{6}\right)$, which sums to 22, – less than the total of 24 attainable by the grand coalition.

To derive the coalition structure $\{\{1\},\{2,3\}\}$, we are implicitly assuming that player 1's decision not to negotiate with 2 and 3 is *irreversible*. Otherwise, once the coalition $\{2,3\}$ has already *formed,* player 1 would have the incentive to join forces with them - - by doing so, he would create an additional surplus of $10\ (24-14),$ of which he could take a bit more than 9 for himself and improve on his payoff from free-riding. But flipflopping on the issue of negotiation like this could have a serious downside: he could acquire a reputation for being a recalcitrant or unserious negotiator. Moreover, if players 2 and 3 anticipate that he will change his mind about negotiating, then it is not clear that the coalition $\{2,3\}$ will even form: if it did form, one of players 2 and 3 would be a payoff of less than 9 even if 1 later joins the coalition (it would be infeasible for all three players to get 9 or more, since 24 is the maximum total payoff). So, *that* player might be better off declining to negotiate, anticipating that the other two players will form a coalition that he can free-ride on.

All of this suggests that player 1 may wish to deliberately *commit* himself not to negotiate. One way he could accomplish this is to cut off communication with players 2 and 3.[19] Imagine, for instance, that initially each player is connected to the others by a telephone line, but that he can later cut the line (thereby severing the means to negotiate with them[20]). Henceforth, we shall be assuming (implicitly or explicitly) that players have the ability to commit themselves not to negotiate.

3. The Model with No Externalities

We consider *n*-player transferable-utility games in characteristic-function form. A *characteristic function v* specifies, for each coalition of players $C \subseteq \{1,\ldots,n\}$, the maximum payoff $v(C)$ that the coalition can achieve. We assume that the game is *super-additive*: if two disjoint coalitions merge, the maximum payoff for the merged coalition is greater than the sum of the individual coalitions' maximum payoffs. Formally, for all disjoint and non-empty coalitions C° and $C^{\circ\circ}$,

(15) $$v(C^{\circ}) + v(C^{\circ\circ}) < v(C^{\circ} \cup C^{\circ\circ})$$ [21]

[19] Thomas Schelling (1961) underscores the usefulness of cutting off communication to prevent further bargaining: "An asymmetry in communications may well favor the one who is unavailable for the receipt of messages, for he cannot be deterred from his own commitment by the receipt of others'" (p.26).

[20] Depending on the application, the telephone line may be more metaphorical than physical. For example, in negotiations between countries, breaking diplomatic ties is one way of avoiding bargaining. Such ties can be restored, but at the cost of political capital.

[21] To assume the weak-inequality version of (15) seems almost without loss of generality. After all, even if two coalitions officially merged, they would presumably still have the option – at worst – of behaving as though they were separate entities. But even the strict-inequality version – which we are adopting – does not appear very demanding. It requires only that when coalitions merge, at least one advantageous strategy that previously was infeasible becomes possible. In any case, we make the assumption partly for rhetorical purposes: it implies that for efficiency, the grand coalition *must* form. Thus, it is a useful counterweight to our argument in this paper that there may be significant forces pushing *against* efficiency.

A *solution concept* $\Phi(\cdot)$ is a function of the characteristic function *v* predicting players' payoff *n*-tuples. Formally, for all *v*,

$$\Phi(v) \subseteq \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \,\middle|\, \sum_{j=1}^{n} x_j \le v(\{1, \ldots, n\}) \right\},$$

i.e., a solution concept picks out a subset of the *n*-tuples that are *feasible* in the sense that the sum of payoffs doesn't exceed the grand coalition's maximum payoff. In particular, the *core* $\Phi^{core}(\cdot)$ is defined to be

(16) $$\Phi^{core}(v) = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \,\middle|\, \sum_{j \in C} x_j \ge v(C) \text{ for all } C, \text{ with equality for } C = \{1, \ldots, n\} \right\},$$ [22]

the set of feasible *n*-tuples that no coalition can block.

Shapley (1953) defines a solution concept (the Shapley value), in which, given *v*, a single *n*-tuple of payoffs $(\varphi_1^*(v), \ldots, \varphi_n^*(v))$ is picked out, where

(16a) $$\varphi_i^*(v) = \sum_{C \subseteq \{1,\ldots,n\}-\{i\}} \frac{|C|!(n-1-|C|-1)!}{n!} (v(C \cup \{i\} - v(C))$$

He shows that this is the unique solution concept that satisfies *Pareto efficiency*, *symmetry* (if *v* treats players symmetrically, then $\varphi_i^*(v)$) is the same for all players), *linearity* $(\varphi^*(v+v') = \varphi(v) + \varphi(v'))$ and *dummy* (if $v(C \cup \{i\}) - v(C) = 0$ for all *C*, then $\varphi_i^*(v) = 0$)

In our approach, the Shapley value is defined or axiomatized as:

[22] In this section (in contrast to Section 4) a solution concept specifics only *payoffs*, and not also which coalitions form. Implicitly, the grand coalition always obtains. Indeed, the assumption of super-additive implies that, for the core, the grand coalition *must* form. If, instead, players were partitioned into disjoint coalitions C° and $C^{\circ\circ}$, then, for feasibility, payoffs would have to sum to $v(C^{\circ}) + v(C^{\circ\circ})$. And so, from (15), the grand coalition could block.

(17) $$\Phi^{sh}(v) = \left\{ (x_1,\ldots,x_n) \in \mathbb{R}^n \,\middle|\, \text{for all } i,\ \sum_{j \in C} x_j = v(C) \text{ for } C = \{i,\ldots,n\} \right\} \text{ (sequentiality).}$$

Formula (17) is an axiom reflecting an implicit sequential bargaining process: if players bargain in the order $1,2,\ldots,n$, then (as suggested in our analysis of Example 1), players *i*'s payoff x_i is determined by what he and subsequent players can get if they form the coalition $\{i,\ldots,n\}$:

$$\sum_{j=i}^{n} x_j = v(i,\ldots,n).$$

Note that Φ^{core} and Φ^{sh} in (16) and (17) are intriguingly similar in form. The only differences are that (i) the weak inequality in (16) is replaced by equality in (17), and (ii) (16) is defined with respect to *all* coalitions *C*, whereas (17) applies only to coalitions of the form $C = \{i, i+1,\ldots,n\}$. Difference (ii) reflects the contrasting implicit processes underlying the core and the Shapley value: in the latter, bargaining proceeds sequentially and so, after a point, only certain coalitions remain relevant (see the discussion of Example 1 in Section 2), whereas, for the former, *all* coalitions remain relevant. Difference (i) also derives from the contrasting bargaining processes. In the case of Shapley, we suppose that what player *i* can get is strongly tied to what the coalition $\{i,\ldots,n\}$ can get; for the core, there are many coalitions that *i* could join, and so the payoff from coalition $\{i,\ldots,n\}$ provides only a lower bound.

Theorem 1: For all *v*, there exists a unique payoff vector $(\varphi_1(v),\ldots,\varphi_n(v)) \in \Phi^{sh}(v)$, where

(18) $$\varphi_i(v) = v(\{i, i+1,\ldots,n\}) - v(\{i+1,\ldots,n\}) \qquad \text{for all } i = 1,\ldots,n$$

Moreover, if the payoffs (18) are averaged over all possible orderings of the players, players obtain their (ordinary) Shapley values as given by (16a).

Proof: Formula (18) follows immediately from (17). As we noted in footnote 16, the Shapley value result follows from Roth (1977). Q.E.D.

4. The General Case

A. *The General Model and the Core*

To allow for externalities, we now make partition functions (rather than characteristic functions) our focus. For every partition $\mathscr{C}$ [23] of $\{1,\ldots,n\}$ and coalition $C \in \mathscr{C}$, a *partition function* $v(\cdot \mid \cdot)$ assigns a maximum payoff $v(C|\mathscr{C})$ to coalition *C* given that the coalition structure $\mathscr{C}$ has formed.

We continue to assume super-additivity, which now implies that, for all partitions $\mathscr{C}$ and all disjoint and nonempty coalitions C° and $C^{\circ\circ} \in \mathscr{C}$,

(19) $$v(C^{\circ}|\mathscr{C}) + v(C^{\circ\circ}|\mathscr{C}) < v(C^{\circ} \cup C^{\circ\circ}|\mathscr{C}^{\cup}),$$

where $\mathscr{C}^{\cup}$ is the same as $\mathscr{C}$ but with C° and $C^{\circ\circ}$ replaced by $C^{\circ} \cup C^{\circ\circ}$.

Since we are no longer presuming that the grand coalition forms automatically, a solution concept must, in general, now predict not just payoffs but coalition structures. We will begin again with the *core*: Formally,

(20) $$\Phi^{core}(v) = \Big\{(\{1,\ldots,n\}, \Big\{(x_1,\ldots,x_n) \in \mathbb{R}^n \Big| \sum_{j \in C} x_j \geq v\big(C \big| C, C_{-}\big), \text{ for all } C \subseteq \{1,\ldots,n\} \text{ with equality for } C = \{1,\ldots,n\}\Big\}\Big\},$$

where $C_{-} = \{1,\ldots,n\} - C$. Note that the grand coalition *is* predicted to form, as it was in the no-externality version of the core. Indeed, (20) is very similar to (16). The only significant difference is that if coalition C breaks off from the grand coalition, it thereby creates the partition C, C_{-}, and so the no-blocking requirement become the inequality in (20).

---

[23] A partition (i.e., a coalition structure) is a collection of coalitions that are mutually exclusive and collectively exhaustive.

B. *The Generalized Shapley Value*

We wish to formulate a generalized Shapley value that arises implicitly from a sequential bargaining process and so predicts not just overall outcomes but also outcomes *contingent* on what has already happened. Thus, if some coalitions involving players $1,\ldots,i-1$ have already formed, the solution concept specifies what coalitions of the remaining players $i,\ldots,n$ arise in response and what those players' payoffs are.

We shall suppose that players form coalition structures that are *sequential*. That is, for players $\{h,h+1,\ldots,k\}$, we obtain a partition of the form $\{h,\ldots,j_2-1\}$, $\{j_2,\ldots,j_3-1\},\ldots,$ $\{j_m,\ldots,k\}$, where $h\leq j_1\leq\ldots\leq j_m\leq k$.[24] Let $\hat{\mathscr{C}}(\cdot)$ be a *contingent partition* mapping and, for all *i*, let $\hat{x}_i(\cdot)$ be a *contingent payoff mapping*. That is, for all *i* and all sequential partitions $\mathscr{C}'$ of $\{1,\ldots,i-1\}$, $\hat{\mathscr{C}}(\mathscr{C}')$ is a sequential partition of $\{i,\ldots,n\}$ contingent on $\mathscr{C}'$, and $\hat{x}_i(\mathscr{C}')$ is player *i*'s Shapley payoff contingent on $\mathscr{C}'$. We also need to consider, contingent on $\mathscr{C}'$, player *i*'s payoff $x_i(\mathscr{C},\{i,\ldots,k\})$ if the not-necessarily-optimal coalition $\{i,\cdots,k\}$ forms. Then, the *generalized Shapley value* is defined as:

$$\Phi^{gsh}(v)=\Big\{(\hat{\mathscr{C}}(\cdot),\big\{(\hat{x}_1(\cdot),\ldots,\hat{x}_n(\cdot))\big|\ \text{for all } i \text{ and all sequential partitions } \mathscr{C}' \text{ of}$$

$$\{1,\ldots,i-1\},$$

$$x_i(\mathscr{C}',\{i,\ldots,k\})+\sum_{j=i+1}^{k}\hat{x}_j(\mathscr{C}',\{i,\ldots,j-1\}) \tag{21}$$

$$=v(\{i,\ldots,k\}\big|\mathscr{C}',\hat{\mathscr{C}}(\mathscr{C}',\{i,\ldots,k\}))\ \text{for all } k\geq i\quad \text{(sequentiality)}$$

[24] Sequentiality is not a restrictive assumption in the important special case of symmetric games (which are analyzed in Section 6).

and

(22) $\hat{x}_i(\mathscr{C}') = x_i(\mathscr{C}', \{i, \ldots, k^*\}) \geq x_i(\mathscr{C}', \{i, \ldots, k\})$ for all $k \geq i$ (optimality),

where

(23) $\{i, \ldots, k^*\} \in \hat{\mathscr{C}}(\mathscr{C}')\}$ (rational expectations)

Axiom (21) is the analog of the sequentiality axiom (17) in the no-externality case. Axiom (22) reflects the fact that, with the GSV, players $i, \ldots, n$ may choose to form a coalition smaller than $\{i, \ldots, n\}$ and will make this choice in order to maximize player *i*'s payoff – whereas the formation of $\{i, \ldots, n\}$ is taken for granted in a setting without externalities. Axiom (23) is the requirement that $\hat{\mathscr{C}}(\mathscr{C}')$ should predict that the optimal coalition $\{i, \ldots, k^*\}$ will form.

Let us discuss this concept in more detail. Imagine that sequential partition $\mathscr{C}'$ of players $1, \ldots, i-1$ has already formed. Even if player *i* ends up joining one of these earlier coalitions, his payoff will be determined (through sequential bargaining) by what he could get by forming a separate coalition. Let $x_i(\mathscr{C}', \{i, \ldots, k\})$ be his payoff if coalition $\{i, \ldots, k\}$ forms. Axiom (22) says that, given $\mathscr{C}'$, *k* is chosen to maximize *i*'s payoff;[25] i.e., *i*'s Shapley payoff is his maximized payoff $\hat{x}_i(\mathscr{C}')$.[26] (Similarly, the Shapley payoff for each other player $j \in \{i, \ldots, k\}$ is $\hat{x}_j(\mathscr{C}', \{i, \ldots, j-1\})$, given that coalitions $\mathscr{C}', \{i, \ldots, j-1\}$ have already formed by the time player *j* arrives). Correspondingly, (23) asserts that, given $\mathscr{C}'$, the payoff-maximizing coalition $\{i, \ldots, k^*\}$ is predicted to form. Axiom (21) says that the sum of the payoffs to players in

[25] In Section 5, where we consider non-cooperative implementations, some implementations have player *i* choosing the coalition, but others have *i* selling that right to subsequent players. In either case, however, player *i* determines the coalition's size by virtue of arriving first.

[26] Of course, to know that $k = k^*$ is payoff-maximizing, we must consider other choices of *k*, which is why $x_i(\mathscr{C}', \{i, \ldots, k\})$ figures in the definition.

$\{i,\ldots,k\}$ should equal the maximum payoff that this coalition can attain, given that $\mathscr{C}'$ has previously formed and that $\hat{\mathscr{C}}(\mathscr{C}',\{i,\ldots,k\})$ is predicted to form subsequently.

Notice that the overall predicted coalition structure $\mathscr{C}^{gsv}$ is $\hat{\mathscr{C}}(\varnothing)$ - - the sequential partition that arises if no previous coalitions have formed.

Let us illustrate the GSV using the game of Example 2. That game is formally defined as

(24)
(a) $v(\{1\}|\{1\},\{2,3\}) = v(\{2\}|\{2\},\{1,3\}) = v(\{3\}|\{3\},\{1,2\}) = 9;$

(b) $v(\{1,2\}|\{3\},\{1,2\}) = 12;\; v(\{1,3\}|\{2\},\{1,3\}) = 13,\; v(\{2,3\}|\{1\},\{2,3\}) = 14$

(c) $v(\{i\}|\{1\},\{2\},\{3\}) = 0;\quad i = 1,2,3$

(d) $v(\{1,2,3\}|\{1,2,3\}) = 24$

The GSV predicts a unique contingent partition mapping $\hat{\mathscr{C}}(\cdot)$, where

(25)
(a) $\hat{\mathscr{C}}(\varnothing) = \{1\},\{2,3\}$

(b) $\hat{\mathscr{C}}(\{1\}) = \{2,3\}$

(c) $\hat{\mathscr{C}}(\{1,2\}) = \{3\}$

(d) $\hat{\mathscr{C}}(\{1\},\{2\}) = \{3\}$

and unique contingent payoff mappings $\hat{x}_i(\cdot)$ and $x_i(\cdot)$ for $i = 1,2,3$, where

(26)
(a) $\hat{x}_1(\varnothing) = 9 \quad x_1(\{1\}) = 9, \quad x_1(\{1,2\}) = -2 \quad x_1(\{1,2,3\}) = 1$

(b) $\hat{x}_2(\{1\}) = 14, \quad x_2(\{1\},\{2\}) = 0 \quad x_2(\{1\},\{2,3\}) = 14$

(c) $\hat{x}_3(\{1\},\{2\}) = 0, \quad \hat{x}_3(\{1,2\}) = 9$

Note that (25c) and (25d) hold automatically, since after $\{1,2\}$ or $\{1\},\{2\}$ have formed, the only possibility for a distinct coalition is $\{3\}$. If $\{1\}$ has formed and player 3 doesn't join 2 in a coalition, then 2 gets 0 (from (24c)); whereas if $\{2,3\}$ forms, then player 3 gets 0 (since his only alternative is to remain alone), leaving player 2 a payoff of 14 (from (24b)). Hence, (25b) holds. Finally, from (26a), player 1 gets payoff -2 if a coalition $\{1,2\}$ forms (since coalition $\{1,2\}$ produces 12, but player 2 gets 14) and payoff 1 from if {1,2,3} forms (since 2 gets 14, 3 gets 9, and (24d) holds). Hence he is better off on his own, and we obtain (25a).

Theorem 2: For generic[27] *v*, there exists a unique solution to (21) - (23). That is, the generalized Shapley value makes a sharp prediction: a unique contingent coalition mapping and a unique contingent payoff *n*-tuple mapping.

Proof: We will argue by backward induction on *i.* In particular, we will show that, for all sequential partitions $\hat{\mathscr{C}}$ of $\{1,\ldots,i-1\}$ and all $k \geq i$

$$x_i(\mathscr{C}',\{i,\ldots,k\}) = v(\{i,\ldots,k\} \,|\, \mathscr{C}',\{i,\ldots,k\}), \hat{\mathscr{C}}(\mathscr{C}',\{i,\ldots,k\})$$

$\pm$ terms involving *v* as a function of players $j > i$

Suppose first that $i = n$. Consider sequential partition $\mathscr{C}^{n-1}$ of $\{1,\ldots,n-1\}$. Then, since there are no players left after *n*, we have

(27) $$\hat{\mathscr{C}}(\mathscr{C}^{n-1}) = \{n\}$$

(28) $$\hat{x}_n(\mathscr{C}^{n-1}) = v(\{n\} \,|\, \{n\}\ ,\ \mathscr{C}^{n-1})$$

---

[27] Generic uniqueness means that, given $v(\cdot \mid \cdot)$, if the payoffs are perturbed by independent draws from a continuous distribution on $[-\varepsilon,\varepsilon]$, where ε is small, then the probability of uniqueness with the perturbed payoffs is 1.

Suppose next that $i = n-1$. Consider sequential partition $\mathscr{C}^{n-2}$ of $\{1,\ldots,n-2\}$. From super-additivity,

$$v(\{n-1\}\big|\mathscr{C}^{n-2},\{n-1\},\{n\}) + v(\{n\}\big|\mathscr{C}^{n-2},\{n-1\},\{n\}) < v(\{n-1,n\}\big|\mathscr{C}^{n-2},\{n-1,n\})$$

and so

(29) $\qquad \hat{\mathscr{C}}(\mathscr{C}^{n-2}) = \{n-1,n\}$

and from (21), (28) and (29),

$$\hat{x}_{n-1}(\mathscr{C}^{n-2}) = v(\{n-1,n\}\big|\mathscr{C}^{n-2},\{n-1,n\}) - v(\{n\}\big|\mathscr{C}^{n-2},\{n-1\},\{n\}).$$

Finally, suppose that the theorem holds for all $i \geq i^* + 1$. Consider sequential partition $\mathscr{C}'$ of $\{1,\ldots,i^*-1\}$. By inductive hypothesis,

(30) $\qquad \hat{\mathscr{C}}(\mathscr{C}',\{i^*,\ldots,k\})$ and $\hat{\mathscr{C}}(\mathscr{C}',\{i^*,\ldots,j-1\})$ are generically unique for all $k \geq i^*$

(31) $\qquad \hat{x}_j(\mathscr{C}',\{i^*,\ldots,j-1\})$ is generically unique for all *j* with $j > i^*$

and

(32) $\qquad \hat{x}_j(\mathscr{C}',\{i^*,\ldots,j-1\}) = v(\{j,\ldots,k_j^*\}\big|\mathscr{C}',\{i^*,\ldots,j-1\},\hat{\mathscr{C}}(\mathscr{C}',\{i^*,\ldots,j-1\}))$

$\pm$ terms involving *v* as a function of players $h > j$,

for all *j* with $i < j \leq k_j^*$,

where $\{j,\ldots,k_j^*\} \in \hat{\mathscr{C}}(\mathscr{C}',\{i^*,\ldots,j-1\})$. Then, for all $k \geq i^*$, (21), (30) - (32) imply that $x_{i^*}(\mathscr{C}',\{i^*,\ldots,k\})$ is unique and has the form

(33) $\qquad x_{i^*}(\mathscr{C}',\{i^*,\ldots,k\}) = v(\{i^*,\ldots,k\}\big|\mathscr{C}',\{i^*,\ldots,k\},\hat{\mathscr{C}}(\mathscr{C}',\{i^*,\ldots,k\}))$

$\pm$ terms involving *v* as a function of players $h > j$

From the right-hand side of (33), there is generically a unique value of k (i.e., $k = k^*$) that maximizes the left-hand side of (33). Hence, generically, $\hat{\mathscr{C}}(\mathscr{C}')$ is unique and satisfies

$$\hat{\mathscr{C}}(\mathscr{C}') = \{i, \ldots, k^*\}, \ \hat{\mathscr{C}}(\mathscr{C}', \{i^*, \ldots, k^*\})$$

Moreover, $\hat{x}_{i^*}(\mathscr{C}')$ is unique and equals the left-hand side of (33) with $k = k^*$ Q.E.D.

C. *Externalities and No Externalities*

A game in partition-function form $v(\cdot \mid \cdot)$ exhibits *no externalities* if the maximum value of any coalition is independent of the other coalitions that form. Formally, there exists a characteristic function $\hat{v}(\cdot)$ such that for all coalitions C and all partitions $\mathscr{C}$ with $C \in \mathscr{C}$,

(34) $$v(C \mid \mathscr{C}) = \hat{v}(C) .$$

We next show that, when there are no externalities, our generalized Shapley value given by (21) - (23) generically reduces to the ordinary Shapley value given by (18).

Theorem 3: When $v(\cdot \mid \cdot)$ exhibits no externalities, then, generically, the generalized Shapley value predicts that, regardless of what coalitions have already formed, the remaining players will form a single coalition and get payoffs given by (18). Formally, for all i and all sequential partitions $\mathscr{C}'$ of $\{1, \ldots, i-1\}$, we have

$$\hat{\mathscr{C}}(\mathscr{C}') = \{i, \ldots, n\}$$

and

$$\hat{x}_i(\mathscr{C}') = \hat{v}(\{i, \ldots, n\}) - \hat{v}(\{i+1, \ldots, n\}) ,$$

where $\hat{v}$ is given by (34).

Proof: As in the proof of Theorem 2, we argue by backward induction on *i*. Following that previous proof exactly, we can establish the theorem for $i = n$ and $i = n - 1$.

Suppose that the theorem holds for all $i \geq i^* + 1$. Consider sequential partition $\mathscr{C}'$ of $\{i, \ldots, i^* - 1\}$. By inductive hypothesis, we have generically:

(35) $$\hat{\mathscr{C}}(\mathscr{C}', \{i^*, \ldots, k\}) = \{k+1, \ldots, n\} \quad \text{for all } k \geq i^*$$

(36) $$\hat{\mathscr{C}}(\mathscr{C}', \{i^*, \ldots, j-1\}) = \{j, \ldots, n\}) \quad \text{for all } j > i^*$$ [28]

and

(37) $$\hat{x}_j(\mathscr{C}', \{i^*, \ldots, j-1\}) = \hat{v}(\{j, \ldots, n\}) - \hat{v}(\{j+1, \ldots, n\}), \text{for all } j > i^*$$

From (21), and (35) - (37)

(38) $$x_{i^*}(\mathscr{C}', \{i^*, \ldots, k\}) = \hat{v}(\{i^*, \ldots, k\}) - \sum_{j=j^*+1}^{k} \hat{x}_j(\mathscr{C}', \{i^*, \ldots, j-1\})$$

$$= \hat{v}(\{i^*, \ldots, k\}) - \hat{v}(\{i^*+1, \ldots, n\}) + \hat{v}\{(k+1, \ldots, n\})$$

Now, from super-additivity,

$$\hat{v}(\{i^*, \ldots, k\}) + \hat{v}(\{k+1, \ldots, n\}) < v(\{i^*, \ldots, n\}) \text{ for all } k \text{ with } i^* \leq k < n$$

and hence

(39) $$\hat{v}(\{i^*, \ldots, k\}) - \hat{v}(\{i^*+1, \ldots, n\}) + \hat{v}(\{k+1, \ldots, n\})$$

$$< \hat{v}(\{i^*, \ldots, n\}) - \hat{v}(\{i^*+1, \ldots, n\}) \text{ for all } k \text{ with } i^* \leq k < n$$

But from (38), the left-hand side of (39) is $x_{i^*}(\mathscr{C}', (\{i^*, \ldots, k\})$ for $i^* < k < n$ and the right-hand side is $x_{i^*}(\mathscr{C}', \{i^*, \ldots, n\})$. We conclude, from (38) and (39) that

$$\hat{\mathscr{C}}(\mathscr{C}') = \{i^*, \ldots, n\}$$

[28] (35) and (36) assert the same thing using different notation (which will be useful later)

and

$$\hat{x}_{i^*}(\mathscr{C}') = \hat{v}(i^*, \ldots, n) - \hat{v}(i^* + 1, \ldots, n) \qquad \text{Q.E.D.}$$

We will say that a game v exhibits *negative externalities* if whenever two disjoint coalitions merge, the maximum payoffs for other coalitions are reduced. Formally, for all partitions $\mathscr{C}$ and all $C^\circ, C^{\circ\circ} \in \mathscr{C}$ such that $C^\circ \cap C^{\circ\circ} = \varnothing$,

$$v(C|\mathscr{C}) > v(C|\mathscr{C}^\cup), \text{ for all } C \in \mathscr{C} - \{C^\circ, C^{\circ\circ}\},$$

where $\mathscr{C}^\cup$ is the same as $\mathscr{C}$ but with C° and $C^{\circ\circ}$ replaced by $C^\circ \cup C^{\circ\circ}$.

Because we assume super-additivity, two coalitions C° and $C^{\circ\circ}$ can do better ceteris paribus by merging than by remaining disjoint. If, in addition, they impose negative externalities on other coalitions, their merger would seem to give additional impetus to those coalitions to join them. Thus, ultimately, we might expect the grand coalition to form.

This informal argument is accurate in the case $n = 3$ (but, as we will see, is not conclusive for $n > 3$).

Theorem 4: If v exhibits negative externalities and $n = 3$, then

$$\mathscr{C}^{gsv} = \{1, 2, 3\},$$

i.e., the grand coalition forms.

Proof:

From (21)

(40) $$\hat{x}_3(\{1\}, \{2\}) = v(\{3\}|\{1\}, \{2\}, \{3\})$$

(41) $$\hat{x}_3(\{1, 2\}) = v(\{3\}|\{1, 2\}, \{3\})$$

(42) $$x_2(\{1\}, \{2\}) = v(\{2\}|\{1\}, \{2\}, \{3\})$$

(43) $$x_2(\{1\},\{2,3\}) = v(\{2,3\}|\{1\},\{2,3\}) - \hat{x}_3(\{1\},\{2\})$$

and so from (40) and (43)

(44) $$x_2(\{1\},\{2,3\}) = v(\{2,3\}|\{1\},\{2,3\}) - v(\{3\}|\{1\},\{2\},\{3\})$$

From super-additivity, the right-hand side of (42) is less than the right-hand side of (44). Hence,

(45) $$\hat{x}_2(\{1\}) = v(\{2\}|\{1\},\{2,3\}) - v\{3\}|\{1\},\{2\},\{3\})$$

$$\hat{\mathscr{C}}(\{1\}) = \{2,3\}$$

From (21), (40), (41) and (45)

(46) $$x_1(\{1\}) = v(\{1\}|\{1\},\{2,3\})$$

(47) $$x_1(\{1,2\}) = v(\{1,2\}|\{1,2\},\{3\}) - \hat{x}_2(\{1\})$$

$$= v(\{1,2\}|\{1,2\},\{3\}) - v(\{2,3\}|\{1\},\{2,3\})$$

$$+v(\{3\}|\{1\},\{2\},\{3\})$$

and

(48) $$x_1(\{1,2,3\}) = v(\{1,2,3\}|\{1,2,3\}) - \hat{x}_2(\{1\}) - \hat{x}_3(\{1,2\})$$

$$= v(\{1,2,3\}|\{1,2,3\}) - v(\{2,3\}|\{1\},\{2,3\})$$

$$+v(\{3\}|\{1\},\{2\},\{3\}) - v(\{3\}|\{1,2\}\{3\})$$

Now, from negative externalities,

$$v(\{3\}|\{1\},\{2\},\{3\}) - v(\{3\}|\{1,2\}\{3\}) > 0$$

Hence the right-hand side of (48) is greater than

(49) $$v(\{1,2,3\}|\{1,2,3\}) - v(\{2,3\}|\{1\}\{2,3\})$$

From super-additivity, (49) is greater than the right-hand side of (46). Thus we conclude

(50) $$x_1(\{1,2,3\}) > x_1(\{1\}).$$

Now from super-additivity,

$$v(\{1,2,3\}|\{1,2,3\}) - v(\{3\}|\{1,2\},\{3\})$$

$$> v(\{1,2\}|\{1,2\},\{3\})$$

and so

(51) $$x_1(\{1,2,3\}) > x_1(\{1,2\}).$$

From (50) and (51), we conclude that

$$\hat{x}_1(\varnothing) = x_1(\{1,2,3\})$$

$$\hat{\mathscr{C}}(\varnothing) = \{1,2,3\}$$

Q.E.D.

To see that Theorem 4 does not extend to $n > 3$, consider:

Example 3: Consider the 4-player game in which

(52) $v(\{i\}|\{1\},\{2\},\{3\},\{4\}) = 0$ for $i = 1,\ldots,$

(53) $v(\{1\}|\{1\},\{2\},\{3,4\}) = -1, v(\{1\}|\{1\},\{2,3,4\}) = -2, v(\{2\}|\{1\},\{2\},\{3,4\}) = -2$

(54) $v(\{4\}|\{1,2\},\{3\},\{4\}) = -6, v(\{4\}|\{1\},\{2,3\},\{4\}) = -6, v(\{4\}|\{1,2,3\},\{4\}) = -7$

(55) $v(\{1,2\}|\{1,2\},\{3,4\}) = -1, v(\{3,4\}|\{1,2\},\{3,4\}) = 1, v(\{3,4\}|\{1\},\{2\},\{3,4\}) = 2$

(56) $v(\{2,3,4\}|\{1\},\{2,3,4\}) = 3$

(57) $v(\{1,2,3,4\}|\{1,2,3,4\}) = 4$

It is straightforward to verify that $v(\cdot \mid \cdot)$ satisfies super-additivity and exhibits negativity externalities.

We claim that

(58) $\hat{\mathscr{C}}(\varnothing) \neq \{1,2,3,4\}$.

To see that (58) holds, note first, from Theorem 4,

(59) $\hat{\mathscr{C}}(\{1\})=\{2,3,4\}$

From (21), (53)-(57) and (59)

(60) (a) $x_1(\{1,2,3,4\}) + \hat{x}_2(\{1\}) + \hat{x}_3(\{1,2\}) + \hat{x}_4(\{1,2,3\}) = v(\{1,2,3,4\}|\{1,2,3,4\}) = 4$

(b) $\hat{x}_2(\{1\}) + \hat{x}_3(\{1\},\{2\}) + \hat{x}_4(\{1\},\{2,3\}) = v(\{2,3,4\}|\{1,\},\{2,3,4\}) = 3$

(c) $\hat{x}_3(\{1,2\}) + \hat{x}_4(\{1,2\},\{3\}) = v(\{3,4\}|\{1,2\},\{3,4\}) = 1$

(d) $\hat{x}_3(\{1\},\{2\}) + \hat{x}_4(\{1\},\{2\},\{3\}) = v(\{3,4\}|\{1\},\{2\},\{3,4\}) = 2$

(e) $\hat{x}_4(\{1,2,3\}) = v(\{4\}|\{1,2,3\},\{4\}) = -7$

(f) $\hat{x}_4(\{1,2\},\{3\}) = v(\{4\}|\{1,2\},\{3\},\{4\}) = -6$

(g) $\hat{x}_4(\{1\},\{2\},\{3\}) = v(\{4\}|\{1\},\{2\},\{3\},\{4\}) = 0$

(h) $\hat{x}_4(\{1\},\{2,3\}) = v(\{4\}|\{1\},\{2,3\},\{4\}) = -6$

(i) $x_1(\{1\}) = v(\{1\}|\{1\},\{2,3,4\}) = -2$

From (60a) - (60h)

(61) $$x_1(\{1,2,3,4\}) = v(\{1,2,3,4\}|\{1,2,3,4\}) - v(\{2,3,4\}|\{1\},\{2,3,4\})$$

$$+\ v(\{3,4\}|\{1,2\},\{3,4\}) - v(\{4\},|\{1\},\{2\},\{3\},\{4\})$$

$$+\ v(\{4\}\,|\{1\},\{2,3\},\{4\}) - v(\{3,4\},|\{1,2\},\{3,4\})$$

$$+\ v(\{4\}\,|\{1,2\},\{3\},\{4\}) - v(\{4\},|\{1,2,3\},\{4\})$$

$$= -3$$

From (60i) and (61)

$$x_1(\{1\}) > x_1(\{1,2,3,4\})$$

and so from (22) and (23)

$$\hat{\mathscr{C}}(\varnothing) \neq \{1,2,3,4\}, \text{ as claimed}$$

5. Non-Cooperative Implementation

Most cooperative game theory solutions are consistent with a substantial variety of non-cooperative games. We shall illustrate this point – which we believe is a significant virtue – for the generalized Shapley value. Specifically, we will present two non-cooperative games whose subgame-perfect equilibrium outcomes coincide with the GSV predictions.

A. Implementation 1

The game consists of *n* periods. In each period *i*, a new player *i* arrives to find the previous players partitioned into a sequential coalition structure $\mathscr{C}' = \{C_1, \ldots, C_m\}$, where $C_1 = \{1, \ldots, j_2 - 1\}$, $C_k = \{j_k, \ldots, j_{k+1} - 1\}$ for $1 < k < m$, and $C_m = \{i_m, \ldots, i-1\}$. Player j_m can choose either to (a) cut off negotiations with *i* and all subsequent players, or to (b) negotiate with *i* by making him an offer b_i.

If j_m chooses (a), then period *i* ends, the coalition structure now becomes $\{C_1, \ldots, C_m, \{i\}\}$, and the game moves to period $i+1$.

If j_m chooses (b), then player *i* either accepts or rejects b_i. If he accepts, then he receives b_i from player j_m, period *i* ends, the coalition structure now becomes $\{C_1,\ldots,C_m \cup \{i\}\}$, and the game moves to period $i+1$. If he rejects, then period *i* ends, the coalition structure now becomes $\{C_1,\ldots,C_m,\{i\}\},$, and the games moves to period $i+1$.

Suppose that the game ends (i.e., period *n* concludes) with coalition structure $\{C_1,\ldots,C_m\}$. Then, each player j_k, $k=1,\ldots,m,$ gets gross payoff $v(C_k \mid C_1,\ldots,C_m)$ (the payoff from his coalition C_k) minus his payments $b_{j_k+1}+\ldots+b_{j_{k+1}-1}$ to other members of the coalition. Each other member *j* of C_k gets b_j.

Theorem 5: Any subgame perfect equilibrium of Implementation I satisfies (21) - (23) and therefore coincides with the GSV.

Proof: As usual, we argue by backwards induction. Suppose first that $i=n$ and $\mathscr{C}'=\{C_1,\ldots,C_m\}$ is a sequential partition of $\{1,\ldots,n-1\}$. Whether or not player *n* joins coalition C_m, his payoff will be $v(\{n\} \mid \mathscr{C}',\{n\})$ (because, even if he does join, player j_m will, in equilibrium, offer him no more than this). Hence, (21) – (23) hold degenerately for $i=n$.

Next, suppose that (21)- (23) hold in any subgame perfect equilibrium for all $i \geq i^*+1$. Consider a sequential partition $\mathscr{C}'=\{C_1,\ldots,C_m\}$ of $\{1,\ldots,i^*-1\}$. If player i^* is to assemble coalition $\{i^*,\ldots,k\}$, then, by inductive hypothesis, she must offer each player $j=i^*+1,\ldots,k$ at least $\hat{x}_j(\mathscr{C}',\{i^*,\ldots,j-1\})$ to get him to join. Moreover, in equilibrium, she won't offer more than this. Hence, from inductive hypothesis, her payoff from assembling $\{i^*,\ldots,k\}$ is

$$v(\{i^*,\ldots,k\} \mid \mathscr{C}',\hat{\mathscr{C}}(\mathscr{C}',\{i,\ldots,k\})) \\ -\sum_{j=i^*+1}^{k}\hat{x}_j(\mathscr{C}',\{i^*,\ldots,j-1\}),$$

and so (21) holds. Because player i^* can terminate coalition-building whenever she wishes by cutting off communication with subsequent players, (22) and (23) will also hold in equilibrium. Q.E.D.

B. Implementation II

The game again consists of *n* periods. In each period *i*, a new player *i* arrives to find the previous players partitioned into a sequential coalition structure $\mathscr{C}' = C_1, ..., C_m$ as in Implementation I. Player $i-1$ (rather than j_m, as in Implementation I) can choose whether to (a) cut off negotiations with *i* and all subsequent players or to (b) negotiate with *i* by making a demand d_i.

If $i-1$ chooses (a), then period *i* ends, the coalition structure now becomes $\{C_1, \ldots, C_m, \{i\}\}$, and the game moves to period $i+1$.

If $i-1$ chooses (b), then player *i* either accepts or rejects d_i. If he accepts d_i, then he pays d_i to player *i*, period *i* ends, the coalition structure now becomes $\{C_1, \ldots, C_m \cup \{i\}\}$, and the game moves to period $i+1$.

Suppose the game ends with coalition structure $\{C_1, \ldots, C_m\}$, where $C_m = \{j_m, \ldots, n\}$. Then, for each $k = 2, \ldots m+1$, player $j_k - 1$ (where $j_{m+1} - 1 = n$) gets gross payoff $v(C_k | C_1, \ldots, C_m)$ from coalition C_k minus the payments $\sum_{j_k}^{j_{k+1}-2} d_j$ she made to the other players in coalition C_k.

Theorem 6: Implementation II implements the GSV in subgame perfect equilibrium.

Notice that the central difference between Implementations I and II is that in I, player j_k (the first member of C_k) buys out the subsequent members of coalition C_k and controls the size of the coalition, whereas in II, player $j_{k+1}-1$ (the last member of C_k) buys out the preceding members of C_k and exercises control. Nevertheless, the two implementations result in the same coalition structure and payoffs.

6. Zipf's Law

Zipf's Law (Zipf 1949) is the empirical observation that the size distribution of many human-created collections (e.g., sets of firms, cities, and words) follows a power law. Specifically, if we examine industries, we find that the number of firms of size x or bigger is approximately c/x, where c is a constant.

There are various explanations of Zipf's Law in the literature, most of them related to firms' growth rates. We will offer an alternative story based on the GSV.

Imagine first that there are three identical firms in an industry. If two of them merge, they will thereby increase their total profit – given superadditivity – but let's assume that they will enhance the profit of the third firm by even more.[29]

We claim that, according to the GSV, there will be one "large" firm (consisting of two merged firms) and one "small" firm (the third firm). This exactly what Zipf's predicts. For $c = 2$

$$2/1 = 2$$
$$2/2 = 1$$

[29] In fact, in a three-firm Cournot model of perfect substitutes, *all* gain in profitability will accrue to the third firm when two firms merge - - the merged firms will themselves earn *less* profit per capita than before (this is called the Cournot merger paradox). Of course, if one relaxes the extreme Cournot assumptions somewhat, mergers can become profitable, but our point is that it is not a strange assumption to posit that the other firms in the industry gain even more than the merged firms.

i.e., there are two firms of size 1 or bigger and one firm of size 2.

We now make this formal. Since we are studying symmetric firms, only the number of players in a coalition is relevant and so we will use the notation

$$w(i|i, j, k)$$

to denote the payoff to a coalition of size *i* if the structure consists of coalitions of size *i*, *j*, and *k*. Let us assume that all externalities are positive.

Theorem 7: Consider a three-player game. Suppose

$$\begin{aligned} w(1|1,2) - w(1|1,1,1) \\ > w(3|3) - (w(1|1,2) + w(2|1,2)) \end{aligned} \tag{62}$$

Then, the GSV predicts a coalition structure in which there is one coalition of size 1 and one of size 2, i.e., Zipf's Law holds.

Remark: The left-hand side of (62) is the external effect on the third player of two firms merging. The right-hand side is the gain that coalitions of sizes 1 and 2 achieve by merging (which is positive from superadditivity).

Proof: To avoid confusing players' names with the coalition sizes, we will rename the players A, B, and C. Let us use Implementation I. If coalition {A, B} has formed, A must pay player C

$$w(1|1,2)$$

to join and form {A, B, C}. Similarly, player A must pay B

$$w(2|1,2) - w(1|1,1,1)$$

to form coalition {A, B}. Thus, A signs up only B, his payoff is

$$w(1|1,2) - [w(2|1,2) - w(1|1,1,1)], \tag{63}$$

which is less than his payoff $w(1|1,2)$ from remaining alone (since the expression in square brackets in (63) is positive from superadditivity). If A signs up both B and C, then his payoff is

(64) $\quad w(3|3)-[w(2|1,2)-w(1|1,1,1)]-w(1|1,2),$

which is less than $w(1|1,2)$ from (62). Q.E.D

Next, suppose that, before mergers, there are *five* identical firms in an industry. Zipf's Law says that three of these firms should merge into one big firm, while the other two remain unmerged, i.e., for $c=3$,

$$3/3=1$$

and

$$3/1=3\,.$$[30]

We will see that the GSV predicts this outcome provided that (i) the additional profit from merging a firm of size 1 with another of size 2 is big enough, and (ii) the external effect on a firm from two others merging (assuming the remaining two are already merged) is sufficiently bigger than the effect in (i).

Let us normalize payoffs so that $w(1|1,1,1,1,1)=0$. Formally, we have:

Theorem 8: Consider a five-player game. Suppose that

(65) $\quad w(3|1,1,3)-(w(1|1,1,1,2)+w(2|1,1,1,2))$ is sufficiently big

and

(66) $\quad w(1|1,2,2)-w(1|1,1,2)$ is sufficiently bigger than the formula in (65) and

$$w(3|2,3)-(w(1|1,2,2)+w(2|1,2,2))\,.$$

Then, the GSV predicts a coalition structure in which there are two coalitions of size 1 and one of size 3, i.e., Zipf's Law holds.

---

[30] $3/2=1.5$, doesn't fit the Law, but this is only an integer problem.

Proof: Let's name the players A, B, C, D, and E.

If A and B are separate, the coalition {C, D} must pay E.

(67) $w(1|1,1,1,2)$

to get him to join the coalition, and player C must pay D

(68) $w(2|1,1,1,2)$

to get D to join C. Hence, given that A and B are separate, C will create the coalition {C, D, E} provided that

$$\begin{aligned} &w(3|1,1,3)-(w(2|1,1,1,2)-w(1|1,1,1,2)) \\ &\quad > w(1|1,1,1,2), \end{aligned} \tag{69}$$

where the left-hand side of (69) is player C's pay off from the three-firm coalition and the right-hand side is his payoff if he remains separate and then coalition {D, E}forms.

But (69) holds from (65), and so coalition {C, D, E} indeed forms if A and B remain separate.

Next, observe that to sign C up, B must pay C the left-hand side of (69), leaving B a payoff of

$$\begin{aligned} &w(2|1,2,2)-(w(3|1,1,3)+w(2|1,1,1,2)) \\ &\quad +(w(1|1,1,1,2), \end{aligned} \tag{70}$$

If, however, B remains by himself, his payoff is

(71) $w(1|1,1,1,3)$

Now, the inequality that (71) is greater than (70) can be written as

$$\begin{aligned} &w(3|1,1,3)-(w(1|1,1,1,2)+w(2|1,1,1,2)) \\ &\quad >(w(2|1,2,2)-w(1|1,1,3), \end{aligned} \tag{72}$$

Which holds from (65).

And so B will remain by himself.

If A signs B up, then we need to show that C will choose to remain separate, in which case his payoff is

(72) $w(1|1,2,2)$

Suppose instead that he signs up D and E. He will need to pay E

(73) $w(1|1,2,2)$

and D

(74) $w(2|1,2,2)-w(1|1,1,1,2)$

Hence C's net payoff will be

(75) $$\begin{aligned} &w(3|2,3)-(w(2|1,2,2)+w(1|1,1,1,2)) \\ &\quad -(w(1|1,2,2) \end{aligned}$$

Now the inequality (72) greater than (75) can be rewritten as

$$\begin{aligned} &w(1|1,2,2)-w(1|1,1,1,1,2) \\ &\qquad > w(3|2,3)-(w(2|1,2,2)+w(1|1,2,2)), \end{aligned}$$

which holds from (66). Hence, C will indeed choose to remain alone.

Finally, we need to show that A will choose to remain by himself and get payoff

(76) $$\begin{aligned} &w(1|1,1,3) \\ &\qquad = w(1|1,1,1,2)+[w(1|1,2,2)-w(1|1,1,1,2)... \\ &\qquad\qquad + w(1|1,1,3)-w(1|1,2,2) \end{aligned}$$

rather than pay B $w(1|1,1,3)$ and achieve a payoff of

(77) $w(2|1,2,2)-w(1|1,1,3)$

But (76) will be greater than (77) provided that

$$
\begin{aligned}
&2[w(1|1,2,2)-w(1|1,1,1,2)]\\
(78)\qquad &\quad > w(2|1,2,2)-2(w(1|1,1,1,2))\\
&\qquad\qquad -2(w(1|1,1,3)-w(1|1,2,2))
\end{aligned}
$$

But (78) holds that to (66). Q.E.D.

We believe that the same arguments can be applied to establish Zipf's Law for larger numbers of firms. For example, with 8 firms Zipf's Law gives us equation

$$
\begin{aligned}
4/1 &= 4\\
4/2 &= 2\\
4/4 &= 1,
\end{aligned}
$$

and a prediction of 1 firm of size 4, 1 firm of size 2, and 2 firms of size 1.

References

Aghion, P., P. Antras, and E. Helpman (2007), "Negotiating Free Trade," *Journal of International Economics* **73**, 1-30.

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

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

Bloch, F. and A. Gomes (2006), "Contracting with Externalities and Outside Options," *Journal of Economic Theory* **127**, 172-201.

Borm, P., Y. Ju, and D. Wettstein (2015), "Rational Bargaining in Games with Coalitional Externalities," *Journal of Economic Theory* **157**, 236-254.

Coase, R. (1960), "The Problem of Social Cost," *Journal of Law and Economics* **3**, 1-44.

de Clippel, G. and R. Serrano (2008a), "Bargaining, Coalitions, and Externalities: A Comment on Maskin," Working Paper 2008-16, Department of Economics, Brown University.

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

Fudenberg, D. And J. Tirole (1991), *Game Theory*, MIT Press, Cambridge, MA.

Gabaix, X. (1999), "Zipf's Law for Cities: An Explanation," *The Quarterly Journal of Economics*, **114**(3), 739-767.

Gillies, D. B. (1959), "solutions to General Non-Zero-Sum Games," in A.W. Tucker and R.D. Luce (eds.) contributions to the *Theory of Games IV*, Princeton University Press, 47-85.

Grabisch, M. and Y. Funaki (2012), "A Coalition Formation Value for Games in Partition Function Form," *European Journal of Operations Research* **221**, 175-185.

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

Hart, S. and A. Mas-Colell (2010), "Bargaining and Cooperation in Strategic Form Games," *Journal of the European Economic Association* **8**, 7-33.

Koczy, L. (2019), *Partition Function Form Games: Coalitional Games with Externalities*, Springer.

Macho-Stadler, I., D. Perez-Castrillo and D. Wettstein (2006), “Efficient Bidding with Externalities,” *Games and Economic Behavior* **57**, 304-320.

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

Macho-Stadler, I., D. Perez-Castrillo and D. Wettstein (2019), “Extensions of the Shapley Value for Environments with Externalities,” In *Handbook of the Shapley Value*, E. Algava et al. (eds.), CRC Press, Taylor and Francis.

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

Maskin, E. (2016), “How Can Cooperative Game Theory Be Made More Relevant to Economics?: An Open Problem,” in J. F. Nash and M. T. Rassias (eds.) *Open Problems in Mathematics*, Springer, 347-350.

McQuillin, B. (2009), “The Extended and Generalized Shapley Value: Simultaneous Consideration of Coalitional Externalities and Coalitional Structure,” *Journal of Economic Theory* **144**, 696-721.

McQuillin, B. and R. Sugden (2016), “Backward Induction Foundations of the Shapley Value,” *Econometrica* **84**, 2265-2280.

Montero, M. (1999), “Coalition Formation in Games with Externalities,” CentER Discussion Paper 121, University of Tilburg.

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

Nash, J. F. (1953), “Two-Person Cooperative Games,” *Econometrica* **21**, 128-140.

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.

Ray, D. and R. Vohra (2014), “Coalition Formation,” in *Handbook of Game Theory* (vol. 4), H. P. Young and S. Zamir (eds.), Elsevier.

Schelling, T. C. (1961) “Experimental Games and Bargaining Theory.” *World Politics*, vol. 14, no. 1, 47-68.

Serrano, R. (2005), “Fifty Years of the Nash Program, 1953-2003,” *Investigaciones Economicas*

29, 219-258.

Serrano, R. (2008), "Nash Program," in *The New Palgrave Dictionary of Economics*, 2nd edition, S. Durlauf and L. Blume (eds.), McMillan, London.

Serrano, R. (2021), "Sixty-Seven years of the Nash Program: Time for Retirement?" *SERIEs, Journal of the Spanish Economic Association* **12**, 35-48.

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

Stigler, G. (1966), *Theory of Price*, University of Chicago Press.

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

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

Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort*. Addison-Wesley Press.