

Bargaining, Coalitions, and Externalities^{*}
E. Maskin
June 2003

1. *Introduction*

For many years now, the cooperative side of game theory has been dominated by the noncooperative side, at least judging from their respective influence on mainstream economics. That cooperative theory should be in relative eclipse is regrettable because this body of work offers us the opportunity to understand how coalitions behave, i.e., how subsets of players bargain over their choice of actions. Such bargaining seems basic to many aspects of economic and political life, from the European Union, to the Kyoto protocol on greenhouse gas emissions, to the OPEC cartel.

Perhaps one reason that cooperative theory has not been more influential on the mainstream is that its two most important solution concepts for games of three or more players, the core and the Shapley value, presume that the *grand* coalition—the coalition of all players—*always* forms. And thus the possibility of interaction *between* coalitions—often important in reality—is ruled out from the beginning.

Indeed, insisting that the grand coalition should always form is highly restrictive, not only realistically but theoretically. As I will argue below, there are many games in which it is implausible to expect that the grand coalition will form. Specifically, these are situations in which coalitions generate significant positive externalities. Roughly speaking, if a player need not join a coalition to derive benefits from its activity, he may have the incentive to remain apart, i.e., to free-ride.

^{*} This is a preliminary and incomplete version of my presidential address to the Econometric Society. I thank the NSF for research support.

In this paper, I develop a cooperative solution concept to handle such situations. The proposed solution, which is derived axiomatically, can be thought of as a generalization of the Shapley value (Shapley, 1953). Just as the Shapley value makes a prediction of the payoffs resulting from bargaining among agents in games in which coalitions do not exert externalities on other coalitions, the generalized solution makes payoff predictions—and also a forecast of which coalitions will form—in games that have such externalities. It reduces to the Shapley value when the externalities disappear.

There is a fairly small but important previous literature on cooperative game theory in which the grand coalition may fail to form, including Aumann and Myerson (1988), Bloch (1996), Greenberg and Weber (1993), Ray and Vohra (1997), (1999), Chwe (1994), Hart and Kurz (1983), and Yi (1997). But to my knowledge, no existing theory gives a fully axiometric treatment of bargaining when externalities constrain efficiency (my own work has been influenced particularly by Ray and Vohra (1999)).

In section 2, I present two examples that motivate my approach. In section 3, I lay out the general model and develop the notion of a solution concept that predicts both payoffs and coalitions. In section 4, I derive the proposed solution concept from a series of four axioms. In section 5, I show that the solution predicts that the grand coalition will form provided that externalities are nonpositive. In section 6, I show that, in the absence of externalities, the predicted payoffs are given by the Shapley value. In section 7, I show that when the core is nonempty, the solution predicts that the grand coalition will form. In section 8, I exhibit a noncooperative implementation of the solution concept. Finally, I offer some suggestions for extensions in section 9.

2. Two examples

There are many games in which the core fails to exist. In such games, any payoffs attainable by the grand coalition can be blocked by some subcoalition. Thus, if there are games in which the grand coalition does not form, one may be tempted to search for them among the class of games without a core. Indeed, I will argue in Theorem 4 below that it is *only* among coreless games that we will find them. However, the fact that the core is empty does not *imply* that the grand coalition will fail to form, as the following example suggests.

Example 1: A Game of Production

Consider a game of three players, a , b , and c . Players a and b together can produce 15, players a and c together can produce 16, and players b and c together can produce 18. Furthermore, the grand coalition can produce 24, and each player on his own can produce 6. All numbers here correspond to payoffs, and utility is transferable

The first thing to notice is that the core of this game is empty. If (x_a, x_b, x_c) is a vector of payoffs, then we must have

$$(1) \quad \begin{aligned} x_a + x_b &\geq 15 \\ x_a + x_c &\geq 16 \\ x_b + x_c &\geq 18 \end{aligned}$$

to prevent the vector from being blockable by a coalition of two players. But adding up the inequalities in (1), we obtain $2(x_a + x_b + x_c) \geq 49$ i.e.,

$$x_a + x_b + x_c \geq \frac{49}{2},$$

which is inconsistent with the fact that the grand coalition can produce at most 24.

Nevertheless, I will argue that it is reasonable to expect the grand coalition to form in this game, provided that agents can sign binding contracts.

To see this, imagine that player a arrives at the bargaining site first, followed by player b , and finally by c . When player b arrives, player a will make him an offer to induce him to form a coalition. What must b be offered to get him to accept? Notice that if b does not accept, then a and b will end up competing for player c . Player a will be willing to offer c a payoff of 16 (the product of a coalition between a and c) minus 6 (what a could get on his own), i.e., 10. Similarly, b will be willing to offer c $18 - 6 = 12$. So, in the bidding for c , b should win and pay (just over) a 's maximum bid of 10. This means that, in the event of a failed negotiation between a and b , b will wind up with a net payoff of $18 - 10 = 8$.

Hence, if b is offered (just over) 8 by player a , he should accept, since 8 is the best he can hope for without a . We would anticipate, therefore, that when b arrives, he and a will sign a binding contract in which b gets a payoff of approximately 8.¹

What will happen when c shows up? His options are either to join the ab coalition or remain on his own, in which case he gets a payoff of 6. Thus, he should be lurable into the coalition for an offer of 6. And since his marginal contribution to that coalition, viz, $24 - 16 = 8$, exceeds 6, the coalition will find it profitable to make the offer.

Thus, despite the emptiness of the core, the grand coalition will indeed form. Indeed, if we suppose that player a receives the residual of the deals with b and c , we can infer that the resulting payoffs will be

¹ This presumes that player a has all the bargaining power. I shall find it convenient to assume that "earlier" players can make take-it-or-leave-it offers to later players. In order to equalize bargaining power we can imagine that order of players' arrivals is chosen randomly.

$$(2) \quad (x_a, x_b, x_c) = (10, 8, 6).$$

Of course, the order of arrival here — a , b , and then c —is accidental. Any of the five other orders are equally possible. But if we perform the corresponding calculations in these other cases, we continue to conclude that the grand coalition will form. Indeed, I will show below (see Theorem 3) that if we average over (2) and its five counterparts, the three players receive exactly their Shapley values $(\frac{43}{6}, \frac{49}{6}, \frac{52}{6})$.

The example illustrates that the concept of the core fails to take into account the sequential nature of bargaining and the binding nature of agreements: At the time that c enters the negotiations, b is already bound to a by contract. And so the fact that b and c could jointly produce 18 is no longer relevant— c will be willing to join coalition ab for considerably less than the productive power of coalition bc would suggest (although that productive power is relevant to b in his earlier negotiations with a).

Example 1 has the property that a coalition's payoff is independent of whatever other coalitions form, a feature shared with most games in the literature on the core and the Shapley value. Notice, for example, that player a can get 6 by himself regardless of whether b and c get together or not. (Indeed, it is this property that validates representing a transferable-utility game in "characteristic function" form, wherein, each possible coalition is identified with its "worth.") Theorem 2 establishes that, as long as this property holds, the grand coalition can be expected to form. However, once the property is relaxed, that expectation may no longer be justified, as the next example illustrates.

*Example 2: A Public Good Game*²

² This game is closely related to Example 1.2 in Ray and Vohra (1999).

Suppose that players a , b , c can produce a public good. The coalition ab can produce 12 for itself, ac can produce 13, and bc can produce 14. The grand coalition can produce 24. A player can produce nothing on his own. However, if the other two players form a coalition, he can free ride on the public good they produce and enjoy a payoff of 9.

I claim that, in this game, we should not expect the grand coalition to form. To see why not, consider the arrival order: a , b , and then c . Let us explore what b must be offered to join a coalition with a . Notice that if he does not join with a , he will be in competition with a for signing up c . In this competition a will be willing to bid 13 (the gross value of the coalition with c) minus 9 (what he would get as a free-rider if b signed up with c), i.e., 4. Similarly, b will be willing to bid $14-9 = 5$. Hence, b will win the bidding war for c and will pay (slightly more than) 4 (notice that because, in this thought experiment, a and b do not form a coalition, c has no possibility of free-riding and so will be willing to accept 4). Hence b 's payoff if he refuses to join with a is $14-4 = 10$. Thus, player a must offer him 10 in order to sign him up.

Assuming b is signed up, player a must then offer player c a payoff of 9 to attract him to coalition ab (because player c has the option to free-ride on ab and get 9 that way). Hence, altogether player a must pay $10+9 = 19$ in order to form the grand coalition. But this leaves only $24-19 = 5$ for himself. Clearly, he would be better off to refrain from signing up b -- in which case, as analyzed above, b will form a coalition with c , and a obtains a free-riding payoff of 9.

I conclude that, with arrival order a , b , c , *two separate* coalitions will form: bc and a by himself. The resulting payoffs are

$$(x_a, x_b, x_c) = (9, 10, 4).$$

The failure of the grand coalition to come about here is because of the externalities. Since players b and c have the option of free-riding (on the externalities created by coalitions ac and ab respectively), player a must offer them each a considerable payoff to lure them into the grand coalition. But this leaves a with so little for himself that in the end he is better off free-riding on *them*. A similar two-coalition outcome obtains for all the other possible arrival orders; the finding that the grand coalition will not form in this model is quite robust. As in Example 1, we can, in fact, calculate the Shapley values (in this case, the “extended” Shapley values) for each player by averaging over these orders. We obtain:

$$(x_a, x_b, x_c) = \left(7, \frac{15}{2}, \frac{49}{6}\right).$$

Notice that these payoffs sum to strictly less than 24, the total attainable by the grand coalition.

The logic behind the no-grand-coalition result depends crucially on externalities being *positive*. By contrast, if, say, the coalition bc exerted a negative externality on player a , this would make a even more eager than without externalities to join with b and c , and so the likelihood of the grand coalition forming would only be *strengthened*. Indeed, Theorem 2 below, establishes that if externalities are negative (or zero), then the grand coalition always forms.

The result also depends on the implicit assumption that a 's decision not to sign up b is irreversible. By this commitment not to merge with b , a in effect forces b and c to form a coalition, and thus gives himself the opportunity to free-ride on them. If instead his decision *could* be changed, players b and c might be reluctant to make a binding coalitional agreement themselves. To see this, note that if they made such an agreement and a was not committed to remain apart, they would then have the incentive to sign him up too (since

this would give them 24 rather than 14). But they would have to pay a at least 9 (his free-riding threat point) to do so, leaving them with only 15 to divide. By contrast, either of them could get a payoff of 9 for himself by sitting back and letting the other two players merge.

We see that if players cannot commit to refrain from forming coalitions, the game develops into a war of attrition in which each player waits for the other two to form a coalition in the expectation of free-riding on them. Such a war gives rise to inefficiencies similar to those created by the commitment not to form coalitions. But since such an avenue is a good deal more complicated than allowing for commitment, I have adopted the latter approach.

3. *The Model*

To capture the possibility of externalities, we consider n -player transferable-utility games in *partition-function form*. More specifically, think of the players $1, \dots, n$ as dividing up into different coalitions. Thus, if each player belongs to exactly one coalition, each possible configuration of coalitions corresponds to a partition \mathcal{C} of $N = \{1, \dots, n\}$. For each partition \mathcal{C} and coalition $C \in \mathcal{C}$, the *partition function* $v(\cdot; \cdot)$ assigns a number $v(C; \mathcal{C})$, interpreted as the payoff to coalition C given the configuration \mathcal{C} . Let us normalize $v(\cdot; \cdot)$ so that $v(\mathbf{f}; \mathcal{C}) = 0$.

I will assume that the partition function $v(\cdot; \cdot)$ is *superadditive*, i.e., that if any *two* coalitions in \mathcal{C} merge, this can only increase their joint payoff (assuming that the other coalitions remain the same):

$$(3) \quad \text{for all } C_1, C_2 \in \mathcal{C} \quad v(C_1; \mathcal{C}) + v(C_2; \mathcal{C}) \leq v(C_1 \cup C_2; \mathcal{C}_2),$$

where \mathcal{C}_{12} is just \mathcal{C} but with C_1 and C_2 replaced by $C_1 \cup C_2$. Behind assumption (3) is the idea that even if two coalitions merge, they always have the option of behaving as they did when they were separate, and so their total payoff should not fall.³

I will be particularly interested in examining what happens to the payoff of a coalition when two *other* coalitions merge. If such a merger has *no* effect, i.e., if for $C, C_1, C_2 \in \mathcal{C}$,

$$(4) \quad v(C; \mathcal{C}) = v(C; \mathcal{C}_{12})$$

(where as before \mathcal{C}_{12} is \mathcal{C} but with C_1 and C_2 replaced by $C_1 \cup C_2$), then I will say that C_1 and C_2 exert *no externality* on C . Games with no externalities can be represented as games in *characteristic function* form. That is, we can express v as a function of C alone. If instead the merger *increases* C 's payoff, i.e.,

$$(5) \quad v(C; \mathcal{C}) < v(C; \mathcal{C}_{12}),$$

then the externality on C is *positive*. Finally, if the payoff *falls*, i.e.,

$$(6) \quad v(C; \mathcal{C}) > v(C; \mathcal{C}_{12}),$$

then the externality is *negative*.

As pointed out already, Example 1 in section 2 is a game with no externalities: the coalition consisting of player a , say, gets a payoff of 6 regardless of whether or not b and c merge. Example 2, by contrast, is a prime example of a game with positive externalities: player a gets nothing on his own if players b and c remain separate, but enjoys a payoff of 9 if they merge.

³ If $v(C; \mathcal{C})$ does not depend on \mathcal{C} (see below), (3) is a mild and standard assumption. It is considerably stronger if the coalition structure affects a coalition's payoff.

For a simple illustration of negative externalities consider the following model:

Example 3: A Pollution Game

Suppose that if players a , b , and c remain on their own they each obtain a payoff of -20 because of the ill effects of pollution. Assume, however, that if any two of them get together, they can partially clean up the pollution and dump the remainder on the third player. In this case the coalition of two gets a payoff of 0 and the third player gets -45 if he is a , -50 if he is b , and -55 if he is c . Finally, if the grand coalition forms, then even more pollution can be cleaned up, resulting in a total payoff of -30 . It follows from Theorem 2 that the grand coalition can be expected to form in this game: there is good reason for at least two players to merge, and once they do, the remaining player has enormous incentive to join too.

Because I am interested in predicting not only payoffs (as the Shapley value and the core do) but the coalitions that form, I shall define a *solution* to consist of a *pair* of functions, one that specifies each player's payoff, the other that specifies the coalitions that form. Examples 1 and 2 adopt the view that coalition formation is a process that occurs sequentially. From this point of view, the order in which players negotiate may matter. Hence, I will index the payoff- and coalition-functions by the *player order* $\mathbf{p} = (k_1, \dots, k_n)$. Notice that there are $n!$ possible orders to which \mathbf{p} can correspond to: the number of permutations of $(1, \dots, n)$. The *overall solution* should then be thought of as a *randomization* over these $n!$ different values that \mathbf{p} can assume.

Given order \mathbf{p} , let $\mathbf{j}^{\mathbf{p}}(\cdot) = (\mathbf{j}_1^{\mathbf{p}}(\cdot), \dots, \mathbf{j}_n^{\mathbf{p}}(\cdot))$ denote a predicted payoff function.

That is, for all i and for v satisfying (3), $\mathbf{j}_i^{\mathbf{p}}(v)$ is a prediction of player i 's payoff if the

order is \mathbf{p} and the partition function is v . Correspondingly, let $\mathbf{y}^p(\cdot)$ denote a predicted coalition function, so that $\mathbf{y}^p(v)$, a partition of N , is the configuration of coalitions predicted to form.

But $(\mathbf{j}^p(\cdot), \mathbf{y}^p(\cdot))$ does not yet fully specify a solution. Given ordering $\mathbf{p} = (k_1, \dots, k_n)$, we should also be able to predict the ultimate payoffs and coalitions conditional on the partial configuration of coalitions that may already have formed. Given $\mathbf{p} = (k_1, \dots, k_n)$, a *partial partition* is a partition of $\{k_1, \dots, k_j\}$ for some $j = 0, 1, \dots, n$ (if $j = 0$, then the partition just consists of the null set). A *solution* is then a pair $(\mathbf{j}^p(\cdot|\cdot), \mathbf{y}^p(\cdot|\cdot))$, where for all v and for all partial partitions \mathcal{E} that have already formed, $\mathbf{j}^p(v|\mathcal{E})$ and $\mathbf{y}^p(v|\mathcal{E})$ are the corresponding predicted payoffs and ultimate coalitions that will follow from \mathcal{E} .

Because, following the discussion of Example 1, I think of the decision to form a coalition as binding, a coalition can be added to but not broken up into pieces (of course, from superadditivity, there would be no advantage to a coalition in breaking up). That is, if $C \in \mathcal{E}$ then there exists $C' \in \mathbf{y}(v|\mathcal{E})$ such that $C \subseteq C'$. Furthermore, following the discussion of Example 2, I think of the decision of player i not to join with j in a coalition as irreversible. Hence, if i and j have joined separate coalitions, these cannot later coalesce. Formally, if $C' \in \mathbf{y}(v|\mathcal{E})$, then there exists at most one $C \in \mathcal{E}$ such that $C \subseteq C'$.

4. *The Axiomatic Characterization*

I now turn to the axioms that characterize the solution. For notational convenience only, I will concentrate on the “natural” order $\mathbf{p}^* = (1, \dots, n)$. Hence, given \mathbf{p}^* , a solution will be denoted as $(\mathbf{j}^{\mathbf{p}^*}(\cdot|\cdot), \mathbf{y}^{\mathbf{p}^*}(\cdot|\cdot))$.

There are four axioms. The first is the requirement of Pareto optimality. Of course, because the grand coalition need not form, we cannot follow Shapley (1953) and require that payoffs sum to the worth of the grand coalition. Instead, we can demand only that, for the configuration of coalitions that *do* form, the payoffs within any coalition sum to its worth:

(i) *Coalitional Pareto Optimality*: For all partial partitions \mathcal{C} , if $C \in \mathbf{y}^{\mathbf{p}^*}(v|\mathcal{C})$, then

$$\sum_{j \in C} \mathbf{j}_j^{\mathbf{p}^*}(v|\mathcal{C}) = v(C; \mathbf{y}(v|\mathcal{C})) \text{ for all } i.$$

The next axiom requires the allocation of players to coalitions to be efficient in a limited sense. Specifically, given the coalitions that have formed among players $1, \dots, i-1$, I shall suppose that player i is allocated to the coalition to which his gross marginal contribution is greatest (where this may be a new coalition, distinct from those that have already formed). Here the gross marginal contribution of player i to a coalition C is the difference between what the members of C get with i (where this figure includes what i gets himself) and what they get if he instead takes the best alternative opportunity. In both computations, what the players “get” is measured by \mathbf{j} , and therefore it takes account of the ultimate configuration of coalitions that will arise.

Formally, given i , let \mathcal{C} be a partition of $\{1, \dots, i-1\}$. For any $C \in \mathcal{C} \cup \mathbf{f}$ and $\hat{C} \in \mathcal{C} \cup \{\mathbf{f}'\}$ ⁴,

let

$$\Phi^i(C, \hat{C}) = \begin{cases} \sum_{j \in C \cup \{i\}} \mathbf{j}^{p_j^*}(v | \mathcal{C}(i, C)), & \text{if } \hat{C} = C \\ \sum_{j \in C} \mathbf{j}^{p_j^*}(v | \mathcal{C}(i, \hat{C})), & \text{if } \hat{C} \neq C, \end{cases}$$

where $\mathcal{C}(i, C)$ is equal to $\mathcal{C} \cup \{\mathbf{f}\}$ but with coalition C replaced by $C \cup \{i\}$. Hence, $\Phi^i(C, \hat{C})$ is the gross payoff to coalition C if i is allocated to coalition \hat{C} , and

$$\Phi^i(C, C) - \Phi^i(C, \hat{C})$$

measures how much bigger the gross payoff to C is when player i joins C rather than \hat{C} . Notice that $\Phi^i(C, C)$ includes the payoff to player i , i.e., it constitutes the *gross* payoff to C (we could alternatively have worked with *net* payoffs).

(ii) *Limited Efficiency*: For all i , all partitions \mathcal{C} of $\{1, \dots, i-1\}$, and all $C^* \in \mathcal{C} \cup \{\mathbf{f}\}$, if

$$(7) \quad i \in C^*, \text{ where } C^* \in \mathbf{y}(v | \mathcal{C}) \text{ and } C^* \subseteq C^*$$

then

$$(8) \quad \begin{aligned} & \Phi^i(C^*, C^*) - \Phi^i(C, C^{**}) \\ & \geq \Phi^i(C, C) - \Phi^i(C, C^*), \text{ for all } C \in \mathcal{C} \cup \{\mathbf{f}\} \text{ such that } C \neq C^*, \end{aligned}$$

where

⁴ $C = \mathbf{f}$ and $\hat{C} = \mathbf{f}'$ correspond to that cases in which C and \hat{C} do not include any of the players $\{1, \dots, i-1\}$. Thus, if i is allocated to coalition $C = \mathbf{f}$, the interpretation is that he has started a new coalition rather than joining any of the existing ones in \mathcal{C} . We distinguish between \mathbf{f} and \mathbf{f}' for technical reasons that will become apparent below.

$$(9) \quad \Phi^i(C^\infty, C^\infty) - \Phi^i(C^\infty, C^\circ) = \max_{\substack{C \neq C^\circ \\ C \in \mathcal{C} \cup \{F\}}} [\Phi^i(C, C) - \Phi^i(C, C^\circ)].^5$$

In words, (ii) says that if, given partition \mathcal{C} of $\{1, \dots, i-1\}$, player i is allocated to coalition $C^\circ \in \mathcal{C} \cup \{F\}$, then i 's gross marginal contribution to C° (the left-hand side of (8)), must be at least as big as his gross marginal contribution to any other coalition C (the right-hand side of (8)).

This kind of efficiency seems plausible because, in the competition among coalitions for player i 's services, the winner should presumably be the one for which i 's gross marginal impact is biggest. Notice, however, that the criterion does not ensure that the allocation of players to coalitions will be fully efficient (indeed, given our assumption of superadditivity, the only fully efficient allocation would result in the grand coalition). This is because in considering whether to assign player i to coalition C° or C , no account is taken on the impact of this assignment on some *third* coalition. Hence, the terminology *limited efficiency*.

The next axiom demands that a player receive his "opportunity wage." That is, his payoff should be his gross marginal contribution to the coalition C^∞ other than C° (the coalition to which he should be allocated) to which he makes the biggest contribution. The

⁵ The maximization defining C^∞ explains why we want to distinguish between \mathbf{f} and \mathbf{f}' : if $C^\circ = \mathbf{f}$, we need to allow for the possibility that C^∞ as well does not belong to \mathcal{C} but is distinct from C° ; formally, this is accomplished by permitting C^∞ to equal \mathbf{f}' rather than \mathbf{f} .

rationale here is that competition for the player's services will drive his compensation up to this level.

(iii) *Opportunity Wages*: For all i and all partitions \mathcal{C} of $\{1, \dots, i-1\}$, if player i is allocated to coalition $C \in \mathcal{C}$, then he receives his opportunity wage, i.e.,

$$\mathbf{j}_i^{p^*}(v|\mathcal{C}) = \Phi^i(C^{\circ\circ}, C^{\circ\circ}) - \Phi^i(C^{\circ\circ}, C^{\circ}),$$

where C° and $C^{\circ\circ}$ satisfy (8) and (9).

The final axiom is just the requirement that \mathbf{j}^{p^*} and \mathbf{y}^{p^*} be consistent in the sense that the payoffs and coalitions that result if the partition \mathcal{C} of $\{1, \dots, i-1\}$ has already formed should be the same as those that result if player i then joins the coalition in $\mathcal{C} \cup \{\mathbf{f}\}$ to which his gross marginal contribution is greatest. More formally, we have:

(iv) *Consistency*: For all i and all partitions \mathcal{C} of $\{1, \dots, i-1\}$, if player i is allocated to coalition $C^{\circ} \in \mathcal{C} \cup \{\mathbf{f}\}$ for which (7) holds,

then

$$\mathbf{j}^{p^*}(v|\mathcal{C}) = \mathbf{j}^{p^*}(v|\mathcal{C}(i, C^{\circ}))$$

and

$$\mathbf{y}^{p^*}(v|\mathcal{C}) = \mathbf{y}^{p^*}(v|\mathcal{C}(i, C^{\circ})).$$

Axioms (i) – (iv) characterize our solution concept:

Theorem 1: For any game v satisfying (3), there exists a solution pair $(\mathbf{j}^{p^*}, \mathbf{y}^{p^*})$ satisfying axioms (i) – (iv). Furthermore, there are only finitely many such solutions.

Remark: Notice that the solution may not be unique. After the proof, I will explain how nonuniqueness arises quite naturally (it is a possible consequence of positive externalities).

Proof: I shall concentrate on the case of three players. The extension to $n > 3$ uses exactly the same methods.

Assume without loss of generality⁶ that

$$(10) \quad v(\{1,3\};\{1,3\},\{2\}) - v(\{1\};\{1\},\{2,3\}) \\ > v(\{2,3\};\{1\},\{2,3\}) - v(\{2\};\{1,3\},\{2\}).$$

Let

$$(11) \quad \begin{cases} \mathbf{j}_i^{P^*} (v|\{1\},\{2\},\{3\}) = v(\{i\};\{1\},\{2\},\{3\}) \text{ for all } i=1,2,3 \\ \mathbf{j}_1^{P^*} (v|\{1\},\{2,3\}) = v(\{1\};\{1\},\{2,3\}) \\ \mathbf{j}_2^{P^*} (v|\{1,3\},\{2\}) = v(\{2\};\{1,3\},\{2\}) \\ \mathbf{j}_3^{P^*} (v|\{1,2\},\{3\}) = v(\{3\};\{1,2\},\{3\}) \end{cases}$$

in accord with axiom (i). There are two cases.

Case I:

$$(12) \quad v(\{2,3\};\{1\},\{2,3\}) - v(\{2\};\{1,3\},\{2\}) \\ > v(\{3\};\{1\},\{2\},\{3\}).$$

Let

$$(13) \quad \mathbf{j}_3^{P^*} (v|\{1\},\{2\}) = \mathbf{j}_3^{P^*} (v|\{1,3\},\{2\}) = \mathbf{j}_3^{P^*} (v|\{1\},\{2,3\}) \\ = v(\{2,3\};\{1\},\{2,3\}) - v(\{2\};\{1,3\},\{2\}).$$

$$(14) \quad \mathbf{j}_2^{P^*} (v|\{1\},\{2\}) = \mathbf{j}_2^{P^*} (v|\{1\},\{2,3\}) = v(\{2\};\{1,3\},\{2\}).$$

$$(15) \quad \mathbf{j}_1^{P^*} (v|\{1\},\{2\}) = \mathbf{j}_1^{P^*} (v|\{1,3\},\{2\}) \\ = v(\{1,3\};\{1,3\},\{2\}) - v(\{2,3\};\{1\},\{2,3\}) + v(\{2\};\{1,3\},\{2\})$$

$$(16) \quad \mathbf{j}_1^{P^*} (v|\{1,2\}) = \mathbf{j}_1^{P^*} (v|\{1,2,3\}) \\ = v(\{1,2,3\};\{1,2,3\}) - v(\{2\};\{1,3\},\{2\}) - v(\{3\};\{1,2\},\{3\})$$

⁶ If inequality (10) is reversed, then whenever player 3 joins 1 in the argument below, have him join player 2 instead..

$$(17) \quad \begin{aligned} \mathbf{j}_2^{P^*}(v|\{1,2\}) &= \mathbf{j}_2^{P^*}(v|\{1,2,3\}) \\ &= v(\{2\};\{1,3\},\{2\}) \end{aligned}$$

$$(18) \quad \begin{aligned} \mathbf{j}_3^{P^*}(v|\{1,2\}) &= \mathbf{j}_3^{P^*}(v|\{1,2,3\}) \\ &= v(\{3\};\{1,2\},\{3\}) \end{aligned}$$

$$(19) \quad \mathbf{y}^{P^*}(v|\{1\},\{2\}) = \{\{1,3\},\{2\}\}$$

$$(20) \quad \mathbf{y}^{P^*}(v|\{1,2\}) = \{\{1,2,3\}\}.$$

We must show that $(\mathbf{j}^{P^*}, \mathbf{y}^{P^*})$ as partially defined by (11) and (13) - (20) satisfies axioms (i) - (iv). Notice first that if we take $\mathcal{C} = (\{1\}, \{2\})$, $C^\circ = \{1\}$, and $C^{\circ\circ} = \{2\}$, then, from (11), (13), and (15),

$$(21) \quad \begin{aligned} \Phi^3(C^\circ, C^\circ) - \Phi^3(C^\circ, C^{\circ\circ}) \\ = v(\{1,3\};\{1,3\},\{2\}) - v(\{1\};\{1\},\{2,3\}); \end{aligned}$$

from (11), (13), and (14),

$$(22) \quad \begin{aligned} \Phi^3(C^{\circ\circ}, C^{\circ\circ}) - \Phi^3(C^{\circ\circ}, C^\circ) \\ = v(\{2,3\};\{1\},\{2,3\}) - v(\{2\};\{1,3\},\{2\}); \end{aligned}$$

and, from (11),

$$(23) \quad \begin{aligned} \Phi^3(\mathbf{f}, \mathbf{f}) - \Phi^3(\mathbf{f}, C^\circ) \\ = v(\{3\};\{1\},\{2\},\{3\}) - 0. \end{aligned}$$

From (10), (12), and (21) – (23), our choices of \mathcal{C} , C° , and $C^{\circ\circ}$ satisfy (8) and (9). Hence, (19) and (20) satisfy axiom (ii); (13) satisfies axioms (iii) and (iv); and (14) and (15) satisfy axioms (i), (iii), and (iv). I conclude that (13) – (20) do indeed satisfy the axioms.

Subcase A:

$$(24) \quad \begin{aligned} &v(\{1,2,3\};\{1,2,3\}) - v(\{2\};\{1,3\},\{2\}) - v(\{3\};\{1,2\},\{3\}) \\ &> v(\{1,3\};\{1,3\},\{2\}) - v(\{2,3\};\{1\},\{2,3\}) + v(\{2\};\{1,3\},\{2\}) \end{aligned}$$

Let

$$(25) \quad \mathbf{j}_i^{p^*}(v) = \mathbf{j}_i^{p^*}(v|\{1\}) = \mathbf{j}_i^{p^*}(v|\{1,2\}), i = 1, 2, 3$$

$$(26) \quad \mathbf{y}^{p^*}(v) = \mathbf{y}^{p^*}(v|\{1\}) = \mathbf{y}^{p^*}(v|\{1,2\}).$$

As before, we must show that $(\mathbf{j}^{p^*}, \mathbf{y}^{p^*})$ as partially defined by (16)-(18), (20), and

(24) - (26) is consistent with the axioms. Now, first take $\mathcal{C} = \{\{1,2\}\}$,

$C^\circ = \{1,2\}$, and $C^\infty = \mathbf{f}$. Then, from (16)-(18) and (25), we have

$$(27) \quad \begin{aligned} &\Phi^3(C^\circ, C^\circ) - \Phi^3(C^\circ, C^\infty) \\ &= v(\{1,2,3\};\{1,2,3\}) - v(\{1,2\};\{1,2\},\{3\}) \end{aligned}$$

and, from (11),

$$(28) \quad \begin{aligned} &\Phi^3(C^\infty, C^\infty) - \Phi^3(C^\infty, C^\circ) \\ &= v(\{3\};\{1,2\},\{3\}) - 0. \end{aligned}$$

Because v is superadditive, (27) exceeds (28), and so our choices of C° and C^∞ satisfy (8) and (9). Hence, (20) and (26) satisfy axioms (ii) and (iv). And (18) and (25) for $i = 3$ satisfy axioms (iii) and (iv).

Next take $\mathcal{C} = \{\{1\}\}$, $C^\circ = \{1\}$ and $C^\infty = \mathbf{f}$. Then, from (15)-(17), and (25) for $i = 1, 2$, we have

$$\begin{aligned}
(29) \quad & \Phi^2(C^\circ, C^\circ) - \Phi^2(C^\circ, C^{\circ\circ}) \\
& = (v(\{1,2,3\}; \{1,2,3\}) - v(\{3\}; \{1,2\}, \{3\})) \\
& \quad - (v(\{1,3\}; \{1,3\}, \{2\}) - v(\{2,3\}; \{1\}, \{2,3\}) + v(\{2\}; \{1,3\}, \{2\})),
\end{aligned}$$

and, from (14),

$$\begin{aligned}
(30) \quad & \Phi^2(C^{\circ\circ}, C^{\circ\circ}) - \Phi^2(C^{\circ\circ}, C^\circ) \\
& = v(\{2\}; \{1,3\}, \{2\}) - 0.
\end{aligned}$$

From (24), (29) exceeds (30), and so these choices of C° and $C^{\circ\circ}$ satisfy (8) and (9).

Hence (20) and (26) satisfy axioms (ii) and (iv); (17) and (25) for $i = 2$ satisfy axioms (iii) and (iv); and (16) and (25) for $i = 1$ satisfy axioms (i) and (iv).

We conclude that, in Case I, Subcase A, formulae (11), (13)–(20), (25), and (26) constitute a solution satisfying all the axioms.

Subcase B:

$$\begin{aligned}
(31) \quad & v(\{1,2,3\}; \{1,2,3\}) - v(\{2\}; \{1,3\}, \{2\}) - v(\{3\}; \{1,2\}, \{3\}) \\
& < v(\{1,3\}; \{1,3\}, \{2\}) - v(\{2,3\}; \{1\}, \{2,3\}) + v(\{2\}; \{1,3\}, \{2\}).
\end{aligned}$$

Let

$$(32) \quad \mathbf{j}_i^{P^*}(v) = \mathbf{j}_i^{P^*}(v | \{1\}) = \mathbf{j}_i^{P^*}(v | \{1\}, \{2\}) \text{ for } i = 1, 2, 3$$

and

$$(33) \quad \mathbf{y}^{P^*}(v) = \mathbf{y}^{P^*}(v | \{1\}) = \mathbf{y}^{P^*}(v | \{1\}, \{2\}).$$

Take $\mathcal{C} = \{\{1\}\}$, $C^\circ = \mathbf{f}$, and $C^{\circ\circ} = \{\mathbf{f}'\}$. Then from (14) and (32),

$$\begin{aligned}
(34) \quad & \Phi^2(C^\circ, C^\circ) - \Phi^2(C^\circ, C^{\circ\circ}) \\
& = v(\{2\}; \{1,3\}, \{2\}) - 0;
\end{aligned}$$

$$(35) \quad \begin{aligned} \Phi^2(C^\circ, C^\circ) - \Phi^2(C^\circ, C^\circ) \\ = v(\{2\}; \{1,3\}, \{2\}) - 0; \end{aligned}$$

and, from (15)-(17), (25) for $i = 1, 2$, and (32),

$$(36) \quad \begin{aligned} \Phi^2(\{1\}, \{1\}) - \Phi^2(\{1\}, C^\circ) \\ = (v(\{1,2,3\}; \{1,2,3\}) - v(\{3\}; \{1,2\}, \{3\})) \\ - (v(\{1,3\}; \{1,3\}, \{2\}) - v(\{2,3\}; \{1\}, \{2,3\}) + v(\{2\}; \{1,3\}, \{2\})). \end{aligned}$$

From (31), (34) exceeds (36), and so these choices of C° and $C^{\circ\circ}$ satisfy (8) and (9). Hence

(33) satisfies axioms (ii) and (iv); and (32) satisfies axioms (i), (iii), and (iv).

We conclude that in Case I, Subcase B, (11), (13)-(20), and (32)-(33) constitute a solution satisfying all the axioms.

Case II:

$$(37) \quad v(\{2,3\}; \{1\}, \{2,3\}) - v(\{2\}; \{1,3\}, \{2\}) < v(\{3\}; \{1\}, \{2\}, \{3\})$$

From superadditivity ((3)) we have

$$(38) \quad v(\{1,3\}; \{1,3\}, \{2\}) - v(\{1\}; \{1\}, \{2\}, \{3\}) > v(\{3\}; \{1\}, \{2\}, \{3\}).$$

Let

$$(39) \quad \begin{aligned} \mathbf{j}_3^{P^*}(v|\{1\}, \{2\}) = \mathbf{j}_3^{P^*}(v|\{1,3\}, \{2\}) = \mathbf{j}_3^{P^*}(v|\{1\}, \{2\}, \{3\}) \\ = v(\{3\}; \{1\}, \{2\}, \{3\}) \end{aligned}$$

$$(40) \quad \mathbf{j}_1^{P^*}(v|\{1\}, \{2\}) = \mathbf{j}_1^{P^*}(v|\{1,3\}, \{2\}) = v(\{1,3\}; \{1,3\}, \{2\}) - v(\{3\}; \{1\}, \{2\}, \{3\})$$

$$(41) \quad \mathbf{j}_2^{P^*}(v|\{1\}, \{2,3\}) = v(\{2,3\}; \{1\}, \{2,3\}) - v(\{3\}; \{1\}, \{2\}, \{3\})$$

$$(42) \quad \mathbf{j}_2^{P^*}(v|\{1\}, \{2\}) = \mathbf{j}_2^{P^*}(v|\{1,3\}, \{2\}) = v(\{2\}; \{1,3\}, \{2\})$$

$$(43) \quad \begin{aligned} \mathbf{j}_1^{P^*}(v|\{1,2\}) &= \mathbf{j}_1^{P^*}(v|\{1,2,3\}) \\ &= v(\{1,2,3\};\{1,2,3\}) - v(\{2\};\{1,3\},\{2\}) - v(\{3\};\{1,2\},\{3\}) \end{aligned}$$

$$(44) \quad \mathbf{j}_2^{P^*}(v|\{1,2\}) = \mathbf{j}_2^{P^*}(v|\{1,2,3\}) = v(\{2\};\{1,3\},\{2\})$$

$$(45) \quad \mathbf{j}_3^{P^*}(v|\{1,2\}) = \mathbf{j}_3^{P^*}(v|\{1,2,3\}) = v(\{3\};\{1,2\},\{3\})$$

$$(46) \quad \mathbf{y}^{P^*}(v|\{1\},\{2\}) = \mathbf{y}^{P^*}(v|\{1\},\{2\}) = \{\{1,3\},\{2\}\}$$

$$(47) \quad \mathbf{y}^{P^*}(v|\{1,2\}) = \{\{1,2,3\}\}.$$

To see that (39)-(47) are consistent with the axioms, first let

$\mathcal{C} = \{\{1\},\{2\}\}$, $C^\circ = \{1\}$, and $C^{\circ\circ} = \mathbf{f}$. Then from (11), (39), and (40),

$$(48) \quad \Phi^3(C^\circ, C^\circ) - \Phi^3(C^\circ, C^{\circ\circ}) = v(\{1,3\};\{1,3\},\{2\}) - v(\{1\};\{1\},\{2\},\{3\});$$

from (39),

$$(49) \quad \Phi^3(C^{\circ\circ}, C^{\circ\circ}) - \Phi^3(C^{\circ\circ}, C^\circ) = v(\{3\};\{1\},\{2\},\{3\}) - 0;$$

and, from (39), (41), and (42)

$$(50) \quad \Phi^3(\{2\},\{2\}) - \Phi^3(\{2\},C^\circ) = v(\{2,3\};\{1\},\{2,3\}) - v(\{2\};\{1,3\},\{2\}).$$

From (37) and (38), and (48) – (50), our choices of C° and $C^{\circ\circ}$ satisfy (8) and (9). Hence

(46) and (47) satisfy axiom (ii); (39) and (42) satisfy axioms (iii) and (iv); and (40) and

(41) satisfy axioms (i) and (iv).

Next, let $\mathcal{C} = \{\{1,2\}\}$, $C^\circ = \{1,2\}$, $C^{\circ\circ} = \mathbf{f}$, then, from (43)-(45)

$$(51) \quad \Phi^3(C^\circ, C^\circ) - \Phi^3(C^\circ, C^{\circ\circ}) = v(\{1,2,3\};\{1,2,3\}) - v(\{1,2\};\{1,2\},\{3\}),$$

and, from (45),

$$(52) \quad \Phi^3(C^{\circ\circ}, C^{\circ\circ}) - \Phi^3(C^{\circ\circ}, C^{\circ}) = v(\{3\}; \{1,2\}, \{3\}) - 0.$$

From (3), (51) exceeds (52), and so our choices of C° and $C^{\circ\circ}$ satisfy (8) and (9). Hence, (47) satisfies axiom (ii) ; and, from (52), (44) satisfies axiom (iii).

Subcase A:

$$(53) \quad v(\{1,2,3\}; \{1,2,3\}) - v(\{2\}; \{1,3\}, \{2\}) - v(\{3\}; \{1,2\}, \{3\}) \\ > v(\{1,3\}; \{1,3\}, \{2\}) - v(\{3\}; \{1\}, \{2\}, \{3\})$$

Assume (39)-(47) and

$$(54) \quad \mathbf{j}_i^{p^*}(v) = \mathbf{j}_i^{p^*}(v|\{1\}) = \mathbf{j}_i^{p^*}(v|\{1,2\}), \quad i=1,2,3$$

and

$$(55) \quad \mathbf{y}^{p^*}(v) = \mathbf{y}^{p^*}(v|\{1\}) = \mathbf{y}^{p^*}(v|\{1,2\}).$$

Take $\mathcal{C} = \{\{1\}\}$, $C^{\circ} = \{1\}$, and $C^{\circ\circ} = \mathbf{f}$. Then, from (40),(43) and (44),

$$(56) \quad \Phi^2(C^{\circ}, C^{\circ}) - \Phi^2(C^{\circ}, C^{\circ\circ}) \\ = (v(\{1,2,3\}; \{1,2,3\}) - v(\{3\}; \{1,2\}, \{3\})) \\ - (v(\{1,3\}; \{1,3\}, \{2\}) - v(\{3\}; \{1\}, \{2\}, \{3\})),$$

and, from (42),

$$(57) \quad \Phi^2(C^{\circ\circ}, C^{\circ\circ}) - \Phi^2(C^{\circ\circ}, C^{\circ}) = v(\{2\}; \{1,3\}, \{2\}) - 0.$$

From (53), (56) exceeds (57), and so our choices of C° and $C^{\circ\circ}$ are consistent with (8) and (9). Thus, (47) and (55) are consistent with axioms (ii) and (iv); (44) and (54) for $i = 2$ are

consistent with axioms (iii) and (iv); and (43) and (54) for $i = 1$ are consistent with axioms (i) and (iv).

We conclude that in Case II, Subcase A, (11), (39)–(47), (54), and (55) constitute a solution satisfying the axioms

Subcase B:

$$(58) \quad \begin{aligned} &v(\{1,2,3\};\{1,2,3\}) - v(\{2\};\{1,3\},\{2\}) - v(\{3\};\{1,2\},\{3\}) \\ &< v(\{1,3\};\{1,3\},\{2\}) - v(\{3\};\{1\},\{2\},\{3\}). \end{aligned}$$

Assume that (39) – (47) and

$$(59) \quad \mathbf{j}_i^{P^*}(v) = \mathbf{j}_i^{P^*}(v|\{1\}) = \mathbf{j}_i^{P^*}(v|\{1\},\{2\}), \quad i=1,2,3$$

and

$$(60) \quad \mathbf{y}^{P^*}(v) = \mathbf{y}^{P^*}(v|\{1\}) = \mathbf{y}^{P^*}(v|\{1\},\{2\}).$$

Take $\mathcal{C} = \{\{1\}\}$, $C^\circ = \mathbf{f}$ and $C^\infty = \{\mathbf{f}^*\}$. Then, from (41)

$$(61) \quad \begin{aligned} &\Phi^2(C^\circ, C^\circ) - \Phi^2(C^\circ, C^\infty) = \Phi^2(C^\infty, C^\infty) - \Phi^2(C^\infty, C^\circ) \\ &= v(\{2\};\{1,3\},\{2\}) - 0, \end{aligned}$$

and from (40), (43) and (44)

$$(62) \quad \begin{aligned} &\Phi^2(\{1\},\{1\}) - \Phi^2(\{1\}, C^\circ) \\ &= (v(\{1,2,3\};\{1,2,3\}) - v(\{3\};\{1,2\},\{3\})) \\ &\quad - (v(\{1,3\};\{1,3\},\{2\}) - v(\{3\};\{1\},\{2\},\{3\})). \end{aligned}$$

From (58), (61) exceeds (62), and so our choices of C° and C^∞ are consistent with (8) and (9). Hence (46) and (60) are consistent with axioms (ii) and (iv); (39), (42), and (59) for $i =$

2,3 are consistent with axioms (iii) and (iv); and (40) and (59) for $i = 1$ are consistent with axioms (i) and (iv).

We conclude that, in Case II, Subcase B, (11), (39) – (47), (59), and (60) constitute a solution satisfying the axioms.

It remains to show that there are only finitely many solutions. To see this, suppose first that $\mathcal{C} = \{\{1\}, \{2\}\}$. Then there are seven possibilities for C° and $C^{\circ\circ}$:

(a) $C^\circ = \{1\}, C^{\circ\circ} = \{2\}$; (b) $C^\circ = \{1\}, C^{\circ\circ} = \mathbf{f}$; (c) $C^\circ = \{2\}, C^{\circ\circ} = \{1\}$; (d) $C^\circ = \{2\}, C^{\circ\circ} = \mathbf{f}$;

(e) $C^\circ = \mathbf{f}, C^{\circ\circ} = \{1\}$; (f) $C^\circ = \mathbf{f}, C^{\circ\circ} = \{2\}$; (g) $C^\circ = \mathbf{f}, C^{\circ\circ} = \mathbf{f}'$. But each of these

possibilities uniquely determines $\mathbf{y}^{p^*}(\nu|\{1\}, \{2\})$. For example, if (b) holds, then axiom

(ii) implies that $\mathbf{y}^{p^*}(\nu|\{1\}, \{2\}) = \{\{1,3\}, \{2\}\}$. But axiom (iii) then determines

$\mathbf{j}_3^{p^*}(\nu|\{1\}, \{2\})$. For example, if again (b) holds, we have

$\mathbf{j}_3^{p^*}(\nu|\{1\}, \{2\}) = \nu(\{3\}; \{1\}, \{2\}, \{3\})$. Moreover, in this case, axiom (i) then determines

that $\mathbf{j}_1^{p^*}(\nu|\{1\}, \{2\}) = \nu(\{1,3\}; \{1,3\}, \{2\}) - \nu(\{3\}; \{1\}, \{2\}, \{3\})$ and

$\mathbf{j}_2^{p^*}(\nu|\{1\}, \{2\}) = \nu(\{2\}; \{1,3\}, \{2\})$.

Similarly, there are only finitely many possibilities for C° and $C^{\circ\circ}$ when

$\mathcal{C} = \{\{1,2\}\}$.

Moving backward, consider $\mathcal{C} = \{\{1\}\}$ next. In this case, there are three possibilities

for C° and $C^{\circ\circ}$: (a) $C^\circ = \{1\}, C^{\circ\circ} = \mathbf{f}$; (b) $C^\circ = \mathbf{f}, C^{\circ\circ} = \{1\}$; (g) $C^\circ = \mathbf{f}, C^{\circ\circ} = \mathbf{f}'$. Now

whichever possibility arises, axiom (ii) and the above determinations of $\mathbf{y}^{p^*}(\nu|\mathcal{C})$ for

$\mathcal{C} = \{\{1\}, \{2\}\}$ and $\mathcal{C} = \{\{1, 2\}\}$ determine $\mathbf{y}^{P^*}(v|\{1\})$ uniquely. For example, if (b) and (b) hold, then

$$(63) \quad \mathbf{y}^{P^*}(v|\{1\}) = \{\{1, 3\}, \{2\}\}.$$

But then in this case, axiom (iv) implies that

$$(65) \quad \mathbf{y}^{P^*}(v) = \{\{1, 3\}, \{2\}\}.$$

Finally, axiom (iv) and (63) imply that

$$(66) \quad \mathbf{j}_i^{P^*}(v) = \mathbf{j}_i^{P^*}(v|\{1\}) = \mathbf{j}_i^{P^*}(v|\{1\}, \{2\}), \quad i = 1, 2, 3.$$

Similarly, a unique pair $(\mathbf{j}_i^{P^*}(\cdot|\cdot), \mathbf{y}^{P^*}(\cdot|\cdot))$ is determined for each of the other $7 \times 3 - 1 = 20$ possibilities.

Q.E.D.

To understand how multiple solutions can arise naturally in games with positive externalities, consider the following example:

Example 4: A bidding war that bidders prefer to lose

The game has 3 players, a, b, c . The partition function is as follows:

$$v(abc; abc) = 28, \quad v(ac; ac, b) = 20, \quad v(b; ac, b) = 8, \quad v(ab; ab, c) = 1,$$

$$v(c; ab, c) = 13, \quad v(bc; a, bc) = 16, \quad v(a; a, bc) = 12,$$

$$v(a; a, b, c) = v(b; a, b, c) = 0, \quad v(c; a, b, c) = 12.$$

Suppose that a and b are competing for player c . Notice that how much a is willing to bid depends on what he anticipates will happen if he loses the competition. If in that case c joins b , then a should be willing to pay $v(ac; ac, b) - v(a; a, bc) = 20 - 12 = 8$. But if

c instead remains alone, then a would pay up to $v(ac; ac, b) - v(a; a, b, c) = 20 - 0 = 20$.

Similarly, player b will pay c $v(bc; bc, a) - v(b; b, ac) = 16 - 8 = 8$, if the alternative is c 's joining a , whereas b will pay up to $v(bc; a, bc) - v(b; a, b, c) = 16 - 0 = 16$, if otherwise c will remain alone.

Observe that player c will require a payment of at least 12 to join either a or b , since he can achieve that payoff on his own. Hence, from the above considerations there are two possible outcomes: either (i) a bids 20, b bids 8, and c joins a for payment of 12 (a bids 20 because $12 > 8$, and so the alternative would be for c to remain alone; b bids 8 because $20 > 12$, and so the alternative would be for c to join a) or else (ii) a bids 8, b bids 16, and c joins b for a payment of 12 (a bids 8 because $16 > 12$, and so the pertinent alternative would be for c to join b ; b bids 16 because $12 > 8$, and so the alternative would be for c to remain alone).

Notice that in case (i), player a wins but ends up with at most $20 - 12 = 8$, which is less than 12, his payoff in case (ii). Similarly, in case (ii), player b wins, but ends up with at most $16 - 12 = 4$, which is less than 8, his payoff in case (i). In other words, each player prefers the outcome in which he loses the bidding war! This is because he does better to free-ride on the agreement made when the other bidder wins than to win himself.

We conclude that there are *two* possible outcomes: in case (i), the ultimate coalition structure will be $\mathbf{y} = \{ac, b\}$ (it can be shown that it would not pay for a and b to get together) and the payoffs $(\mathbf{j}_a, \mathbf{j}_b, \mathbf{j}_c) = (8, 8, 12)$. In case (ii), the coalitions will be $\mathbf{y} = \{a, bc\}$ (again, it can be shown that it does not pay a to get b to merge) and payoffs $(\mathbf{j}_a, \mathbf{j}_b, \mathbf{j}_c) = (12, 4, 12)$.

5. Nonpositive Externalities

Example 4 relies critically on externalities being positive. If instead all externalities are nonpositive, then, as I will now show, the solution is unique and the grand coalition forms.

Theorem 2: Consider a game v satisfying (3) in which all externalities are nonpositive.

There exists a unique pair $(\mathbf{j}^{p^*}, \mathbf{y}^{p^*})$ satisfying axioms (i) – (iv). Furthermore, $\mathbf{y}(v) = N$.

Proof: Once again, I will confine attention, for the sake of simplicity, to the case $n = 3$. As in the proof of Theorem 1, we can assume, without loss of generality, that (10) holds.

I first claim that, in any solution, we must have:

$$(67) \quad \text{if } \mathcal{C} = \{\{1,2\}\}, \text{ then } C^\circ = \{1,2\} \text{ and } C^{\circ\circ} = \mathbf{f}.$$

From the proof of Theorem 1, we know that superadditivity implies there exists a solution satisfying (67). It remains to show that (67) holds in *any* solution.

Suppose, to the contrary, that

$$(68) \quad C^\circ = \mathbf{f}.$$

Then,

$$(69) \quad \Phi^3(C^\circ, C^\circ) - \Phi^3(C^\circ, C^{\circ\circ}) = v(\{3\}; \{1,2\}, \{3\}) - 0$$

and

$$(70) \quad \begin{aligned} & \Phi^3(\{1,2\}, \{1,2\}) - \Phi^3(\{1,2\}, C^\circ) \\ & = v(\{1,2,3\}; \{1,2,3\}) - v(\{1,2\}; \{1,2\}, \{3\}). \end{aligned}$$

But (3) implies that (70) exceeds (69), a contradiction of (8) in axiom (ii). Hence, (67) must hold after all.

Suppose next that $\mathcal{C} = \{\{1\}, \{2\}\}$. I claim that, in any solution,

$$(71) \quad \text{if } \mathcal{C} = \{\{1\}, \{2\}\}, \text{ then } C^\circ = \{1\} \text{ and } C^\infty = \{2\}.$$

If instead

$$(72) \quad C^\circ = \{1\} \text{ and } C^\infty = \mathbf{f},$$

then

$$(73) \quad \Phi^3(C^\infty, C^\infty) - \Phi^3(C^\infty, C^\circ) = v(\{3\}; \{1\}, \{2\}, \{3\})$$

and

$$(74) \quad \begin{aligned} \Phi^3(\{2\}, \{2\}) - \Phi^3(\{2\}, C^\circ) \\ = v(\{2,3\}; \{1\}, \{2,3\}) - v(\{2\}; \{1,3\}, \{2\}). \end{aligned}$$

But because externalities are nonpositive,

$$(75) \quad v(\{2\}; \{1,3\}, \{2\}) \leq v(\{2\}; \{1\}, \{2\}, \{3\}),$$

and so from (3) and (75), (74) exceeds (73), a contradiction of (9). Hence, (72) cannot

hold. If

$$(76) \quad C^\circ = \{2\},$$

then

$$(77) \quad \begin{aligned} \Phi^3(C^\circ, C^\circ) - \Phi^3(C^\circ, C^\infty) \\ = v(\{2,3\}; \{1\}, \{2,3\}) - v(\{2\}; \{1\}, \{2\}, \{3\}), \text{ if } C^\infty = \mathbf{f} \end{aligned}$$

or

$$= v(\{2,3\}; \{1\}, \{2,3\}) - v(\{2\}; \{1,3\}, \{2\}), \text{ if } C^\infty = \{1\}.$$

and

$$(78) \quad \Phi^3(\{1\}, \{1\}) - \Phi^3(\{1\}, C^\circ) \\ = v(\{1,3\}; \{1,3\}, \{2\}) - v(\{1\}; \{1\}, \{2,3\}).$$

From (77) and because externalities are nonpositive,

$$(79) \quad \Phi^3(C^\circ, C^\circ) - \Phi^3(C^\circ, C^{\circ\circ}) \leq v(\{2,3\}; \{1\}, \{2,3\}) - v(\{2\}; \{1,3\}, \{2\}).$$

But from (10), the right-hand side of (78) exceeds that of (79), a contradiction of (8). We conclude that (76) cannot hold. The last remaining possibility is

$$(80) \quad C^\circ = \mathbf{f},$$

in which case we have

$$(81) \quad \Phi^3(C^\circ, C^\circ) - \Phi^3(C^\circ, C^{\circ\circ}) = v(\{3\}; \{1\}, \{2\}, \{3\}) - 0$$

and

$$(82) \quad \Phi^3(\{1\}, \{1\}) - \Phi^3(\{1\}, C^\circ) = v(\{1,3\}; \{1,3\}, \{2\}) - v(\{1\}; \{1\}, \{2\}, \{3\}).$$

But superadditivity implies that (82) exceeds (81), a contradiction of (8). Hence, (71) holds as claimed.

Finally, consider $\mathcal{E} = \{\{1\}\}$. I claim that

$$(83) \quad \text{if } \mathcal{E} = \{\{1\}\}, \text{ then } C^\circ = \{1\} \text{ and } C^{\circ\circ} = \mathbf{f}.$$

To see this, note first that, if $C^\circ = \{1\}$ and $C^{\circ\circ} = \mathbf{f}$, then, from (67) and (71),

$$(84) \quad \Phi^2(C^\circ, C^\circ) - \Phi^2(C^\circ, C^{\circ\circ}) \\ = (v(\{1,2,3\}; \{1,2,3\}) - v(\{3\}; \{1,2\}, \{3\})) \\ - ((v(\{1,3\}; \{1,3\}, \{2\}) - v(\{2,3\}; \{1\}, \{2,3\})) + v(\{2\}; \{1,3\}, \{2\}));$$

and, from (71),

$$(85) \quad \Phi^2(C^\circ, C^\circ) - \Phi^2(C^\circ, C^\circ) = v(\{2\}; \{1,3\}, \{2\}) - 0.$$

But, from (3),

$$(86) \quad v(\{1,2,3\}; \{1,2,3\}) > v(\{1,3\}; \{1,3\}, \{2\}) + v(\{2\}; \{1,3\}, \{2\}).$$

Furthermore, because externalities are nonpositive, we have (75) and

$$(87) \quad v(\{3\}; \{1,2\}, \{3\}) \leq v(\{3\}; \{1\}, \{2\}, \{3\}),$$

implying, from (3), that

$$(88) \quad v(\{2,3\}; \{1\}, \{2,3\}) > v(\{3\}; \{1,2\}, \{3\}) + v(\{2\}; \{1,3\}, \{2\}).$$

Now (86) and (88) imply that (84) exceeds (85), and so this choice of C° and C° satisfies

(8) and (9).

We next must show that we cannot have

$$(89) \quad C^\circ = \mathbf{f}.$$

But if (89) holds, then, from (71),

$$(90) \quad \Phi^2(C^\circ, C^\circ) - \Phi^2(C^\circ, C^\circ) = v(\{2\}; \{1,3\}, \{2\}) - 0.$$

and

$$(91) \quad \begin{aligned} & \Phi^2(\{1\}, \{1\}) - \Phi^2(\{1\}, C^\circ) \\ &= v(\{1,2,3\}; \{1,2,3\}) - v(\{3\}; \{1,2\}, \{3\}) \\ & \quad - (v(\{1,3\}; \{1,3\}, \{2\}) - v(\{2,3\}; \{1\}, \{2,3\}) + v(\{2\}; \{1,3\}, \{2\})). \end{aligned}$$

But, as argued above, (91) exceeds (90), thus violating (8). We conclude, therefore, that

(89) is impossible and that (83) holds.

Together (67), (71), (83) imply that $(\mathbf{j}^{P^*}, \mathbf{y}^{P^*})$ are unique. Furthermore (83) and axioms (ii) and (iv) imply that if $2 \in C$, where $C \in \mathbf{y}^{P^*}(v)$ then $1 \in C$. And (67) and axiom (ii) imply that $\mathbf{y}^{P^*}(v | \{1,2\}) = \{1,2,3\}$. Hence the grand coalition forms.

Q.E.D.

6. Connection with the Shapley Value

I stated at the outset that the solution $(\mathbf{j}^P, \mathbf{y}^P)$ is a generalization of Shapley value in that it reduces to the ordinary value in the case of no externalities. Let us make this formal:

Theorem 3: Consider a game v satisfying (3) for which there are no externalities. Then, if $(\mathbf{j}^P, \mathbf{y}^P)$ satisfies axioms (i)-(iv), it predicts that the grand coalition will form, and, for all i ,

$$(92) \quad \frac{1}{n!} \sum_P \mathbf{j}_i^P(v) = \text{Shapely value for } i \text{ in game } v.$$

Proof: That the grand coalition will form when there are no externalities follows from Theorem 2. To establish (92), we will use the well-known fact that a player's Shapley value is his expected marginal product when players arrive and join the grand coalition in random order. Because we are assuming no externalities, we can express the worth of a coalition without reference to the partition of coalitions that have formed, i.e., for coalition C , we can write $v(C)$. Once again, we will concentrate on the case $n=3$.

Notice first that for the ordering $\mathbf{p}^{231} = (2,3,1)$ or $\mathbf{p}^{321} = (3,2,1)$, player 1's opportunity wage is just the payoff that he can obtain on his own, the alternative to joining the coalition $\{2,3\}$. Hence, from axiom (iii),

$$(93) \quad \mathbf{j}_1^{\mathbf{p}^{231}}(v) = \mathbf{j}_1^{\mathbf{p}^{321}}(v) = v(\{1\}).$$

Case I

$$(94) \quad v(\{2,3\}) - v(\{2\}) > v(\{1,3\}) - v(\{1\})$$

In this case, (94) and axioms (ii), (iii) and (iv) imply that

$$(95) \quad \mathbf{j}_2^{\mathbf{p}^*}(v) = v(\{2,3\}) - v(\{1,3\}) + v(\{1\}).$$

From the proof of Theorem 1,

$$(96) \quad \mathbf{j}_3^{\mathbf{p}^*}(v) = v(\{3\}).$$

Hence, from axiom (i), (95) and (96),

$$(97) \quad \mathbf{j}_1^{\mathbf{p}^*}(v) = v(\{1,2,3\}) - v(\{2,3\}) + v(\{1,3\}) - v(\{1\}) - v(\{3\}).$$

Also, from (94), we have

$$(98) \quad \mathbf{j}_1^{\mathbf{p}^{213}}(v) = v(\{1\}).$$

Next consider the ordering $\mathbf{p}^{132} = (1,3,2)$.

Subcase A

$$(99) \quad v(\{2,3\}) - v(\{3\}) > v(\{1,2\}) - v(\{1\}).$$

Given (99), we obtain

$$(100) \quad \mathbf{j}_3^{\mathbf{p}^{132}}(v) = v(\{2,3\}) - v(\{1,2\}) + v(\{1\})$$

$$(101) \quad \mathbf{j}_2^{\mathbf{p}^{132}}(v) = v(\{2\}),$$

and, so

$$(102) \quad \mathbf{j}_1^{p^{132}} = v(\{1,2,3\}) - v(\{2,3\}) + v(\{1,2\}) - v(\{1\}) - v(\{2\}).$$

Similarly, for $\mathbf{p}^{312} = (3,1,2)$, we obtain, from (99),

$$(103) \quad \mathbf{j}_1^{p^{312}} = v(\{1\}).$$

Combining (93), (97), (98), (102) and (103) we obtain

$$(104) \quad \frac{1}{6} \sum_p \mathbf{j}_i^p(v) = \frac{2}{3} (v(\{1,2,3\}) - v(\{2,3\})) + \frac{1}{3} (v(\{1,3\}) - v(\{1\})) \\ + \frac{1}{3} (v(\{1,2\}) - v(\{2\})) + \frac{2}{3} v(\{1\}).$$

But the right-hand side of (104) is the Shapley value for player 1, as claimed.

Subcase B

$$(105) \quad v(\{2,3\}) - v(\{3\}) < v(\{1,2\}) - v(\{1\})$$

Given (105) we obtain (101) and

$$(106) \quad \mathbf{j}_3^{p^{132}}(v) = v(\{3\})$$

and so

$$(107) \quad \mathbf{j}_1^{p^{132}} = v(\{1,2,3\}) - v(\{2\}) - v(\{3\}).$$

From (105),

$$(108) \quad \mathbf{j}_1^{p^{312}} = v(\{1,2\}) - v(\{2,3\}) + v(\{3\}).$$

Summing (93), (97), (98), (107), (108), we again obtain (104).

Case II

$$(109) \quad v(\{2,3\}) - v(\{2\}) < v(\{1,3\}) - v(\{1\})$$

In this case, we have

$$(110) \quad \mathbf{j}_2^{p^*}(v) = v(\{2\}).$$

From (96) and (110) we have

$$(111) \quad \mathbf{j}_1^{p^*}(v) = v(\{1,2,3\}) - v(\{2\}) - v(\{3\}).$$

Also from (109),

$$(112) \quad \mathbf{j}_1^{p^{213}} = v(\{1,3\}) - v(\{2,3\}) + v(\{2\}).$$

Subcase A: (99) holds

Given (99) we obtain (102) and (103). Summing (93), (111), (112), (102), and (103), we again obtain (104).

Subcase B: (105) holds

Given (105), we obtain (107) and (108). Adding up (93), (111), (112), (107), and (108), we once again obtain (104).

Q.E.D.

7. Connection with the Core

We noted in section 2 that, just because the core is empty, we cannot conclude that grand coalition fails to form. However, the converse *does* hold: if the core exists, the grand coalition must form.

We shall define the *core* of a game v to be the set of payoffs that are unblockable by any coalition C formally,

$$\text{core}(v) = \left\{ (x_1, \dots, x_n) \mid \text{for all coalitions } C, \sum_{i \in C} x_i \geq v(C; \{C, N \setminus C\}) \right\}.$$

This definition presumes that to evaluate the worth of coalition C we should assume that it faces the complementary coalition $N \setminus C$.

Theorem 4: Consider a game v satisfying (3). If the core is nonempty, then, for all \mathbf{p} ,

$$\mathbf{y}^{\mathbf{p}}(v) = N.$$

Proof: I will argue, as before, for the case $n = 3$. Suppose, without loss of generality, that (10) holds.

Consider the order $\mathbf{p}^{213} = (2, 1, 3)$. If the grand coalition fails to form, then,

following the argument in the proof of theorem 1, we obtain

$$(113) \quad v(\{1, 2, 3\}; \{1, 2, 3\}) - v(\{1\}; \{1\}, \{2, 3\}) - v(\{3\}; \{1, 2\}, \{3\}) < v(\{2\}; \{1, 3\}, \{2\}).$$

We must show that the core is empty. Suppose, to the contrary, that the payoff vector

(x_1, x_2, x_3) belongs to the core. Then

$$(114) \quad x_1 + x_2 + x_3 = v(\{1, 2, 3\}; \{1, 2, 3\})$$

$$(115) \quad x_1 \geq v(\{1\}; \{1\}, \{2, 3\}),$$

$$(116) \quad x_2 \geq v(\{2\}; \{1, 3\}, \{2\}),$$

$$(117) \quad x_3 \geq v(\{3\}; \{1, 2\}, \{3\}),$$

where (114) asserts that (x_1, x_2, x_3) should not be blockable by the grand coalition, and

(115), (116), and (117) assert, respectively that it should not be blockable by the singleton coalitions $\{1\}$, $\{2\}$, and $\{3\}$. But from (114)-(117), we obtain

$$\begin{aligned} & v(\{1\}; \{1\}, \{2, 3\}) + v(\{2\}; \{1, 3\}, \{2\}) + v(\{3\}; \{1, 3\}, \{2\}) \\ & \leq v(\{1, 2, 3\}; \{1, 2, 3\}), \end{aligned}$$

which contradicts (113).

Q.E.D.

8. Noncooperative Implementation

There is a variety of noncooperative games that implement the generalized Shapley value. The following is a particularly simple game form that implements $(\mathbf{j}^{p^*}, \mathbf{y}^{p^*})$ in perfect equilibrium.

For $i = 1, \dots, n-1$, let \mathcal{C} be a partition of $\{1, \dots, i-1\}$. For each $C \in \mathcal{C}$ let j_C be the player with the lowest index in C . If partition \mathcal{C} has already formed, then simultaneously each j_C makes a nonnegative bid $b_{j_C}^i$ for player i to join coalition C . Player i then has the option of accepting one of these bids, in which case he joins C and obtains payoff $b_{j_C}^i$, or else he rejects all the bids and starts his own coalition. If, after each player $j = 2, \dots, n$ has decided which coalition to join, the resulting configuration of coalitions is \mathcal{C}' (a partition of $\{1, \dots, n\}$), then, for each $C \in \mathcal{C}'$, j_C 's payoff is

$$v(C; \mathcal{C}') - \sum_{\substack{i \in \mathcal{C}' \\ i \in j_C}} b_{j_C}^i.$$

Theorem 5: Given order \mathbf{p}^* , then, if $(\mathbf{j}^{p^*}, \mathbf{y}^{p^*})$ is a solution (i.e., it satisfies axioms (i)-(iv)), there exists a perfect equilibrium of the above game form for which the equilibrium payoffs and coalitions are given by $(\mathbf{j}^{p^*}, \mathbf{y}^{p^*})$. Furthermore, any equilibrium of the game form corresponds to a solution.

9. Extensions

There are several directions in which I think it would be desirable to extend the analysis.

The first is to allow for the possibility of a player's having *multiple memberships*. The partition function form presumes that coalitions are mutually exclusive, but in reality, a player might belong to multiple coalitions that interact with one another (e.g., a country might belong to both the United Nations and the European Union).

The second is to introduce the possibility of *networks*. In the current analysis, forming coalitions is a transitive relation: if a has formed a coalition with b and b has done so with a , then it is presumed that a has formed a coalition with c , i.e., the coalition consists of a, b, c . But if we think of coalitions as trading relationships, for example, then there is no reason for transitivity; just because a trades with b and b trades with c , we cannot conclude that a trades with c . Following Jackson and Wolinsky (1996), we can define a network as a nondirected graph, in which a node corresponds to a player and a branch between two nodes corresponds to a coalitional relationship between those two players.

Third, it would clearly be desirable to be able to generalize the analysis to *nontransferable utility*. For the reasons discussed in Gul (1989), however, it might be difficult to establish a connection with the (NTU) Shapley value.

Finally, I have presumed throughout that the partition function $v(\cdot|\cdot)$ and all agreements among players are common knowledge. It would certainly be of interest to relax this supposition and allow for *incomplete information*.

References (Incomplete)

- Aumann, R. and R. Myerson (1998), "Endogenous Formation of Links between Players and of Coalitions," in A. Roth (ed.), *The Shapley Value*, Cambridge: Cambridge University Press, 175-191.
- Bloch, F. (1996), "Sequential Formation of Coalitions with Fixed Payoff Division," *Games and Economic Behavior*, 14: 90-123.
- Chwe, M. (1994), "Farsighted Coalitional Stability," *Journal of Economic Theory*, 63: 299-325.
- Greenberg, J. and S. Weber (1993), "Stable Coalition Structures with a Unidimensional Set of Alternatives," *Journal of Economic Theory*, 60: 62-82.
- Gul, F. (1989), "Bargaining Foundations of Shapley Value," *Econometrica*, 57:81-95.
- Hart, S. and M. Kurz (1983), "Endogenous Formation of Coalitions," *Econometrica*, 52: 1047-1064.
- Jackson, M. and A. Wolinsky (1996), "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71: 44-74.
- Ray, D. and R. Vohra (1997), "Equilibrium Binding Agreements," *Journal of Economic Theory*, 73: 30-78.
- Ray, D. and R. Vohra (1999), "A Theory of Endogenous Coalition Structures," *Games and Economic Behavior*, 26: 286-336.
- Shapley, L. (1953), "A Value for n -Person Games," in H. Kuhn and A. Tucker (eds.), *Contributions to the Theory of Games, II*, Princeton: Princeton University Press, 307-317.

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