Games and Economic Behavior **26**, 286–336 (1999)
Article ID game.1998.0648, available online at http://www.idealibrary.com on IDEAL®

A Theory of Endogenous Coalition Structures*

Debraj Ray

Department of Economics, Boston University, Boston, Massachusetts 02215 and Instituto de Análisis Económico (CSIC), 08193 Bellaterra, Barcelona, Spain

and

Rajiv Vohra†

Department of Economics, Brown University, Providence, Rhode Island 02912

Received October 2, 1996

Consider an environment with widespread externalities, and suppose that binding agreements can be written. We study coalition formation in such a setting. Our analysis proceeds by defining on a partition function an extensive-form bargaining game. We establish the existence of a stationary subgame perfect equilibrium for such a game. Our main results are concerned with the characterization of equilibrium *coalition structures*. We develop an algorithm that generates (under certain conditions) an equilibrium coalition structure. Our characterization results are especially sharp for *symmetric* partition functions. In particular, we provide a uniqueness theorem and apply our results to a Cournot oligopoly. *Journal of Economic Literature* Classification Numbers: C71, C72, C78, D62. © 1999 Academic Press

1. INTRODUCTION

1.1. *Background*

We study endogenous coalition formation in contexts where individual (and group) payoffs depend on the entire coalition structure that might

* Preliminary versions of some results contained in this paper were circulated as "Binding Agreements and Coalitional Bargaining" at a workshop on "Endogenous Coalition Formation" that we conducted at Stanford University, July 1995. We thank Paul Milgrom and the participants of that workshop for many valuable comments that have influenced the writing of the current version. We gratefully acknowledge support under National Science Foundation Grants SBR-9414114 (Ray) and SBR-9414142 (Vohra). Partial assistance under Grant PB90-0172 from the Ministerio de Educación y Ciencia, Government of Spain (Ray), and a Fulbright Research Award (Vohra) also supported this research.

† E-mail: rajiv_vohra@brown.edu.

0899-8256/99 $30.00

form. A broader objective is to take a step toward the understanding of coalitional influence in the negotiation process.

Of course, cooperative game theory has been much concerned with this problem. But the major part of this theory is based on the *characteristic function*, which by its very construction assumes away the interesting strategic interactions.

The standard recipe for generating characteristic functions is a minimax argument: if a coalition wishes to go off on its own, it is then presumed to fear the worst, namely, that other coalitions will act in such a way as to minimize the payoffs of the deviant group. This argument creates a set of payoffs for each coalition, and therefore a characteristic function.[1]

Given the amount of energy that has been expended on cooperative game theory from the characteristic function onwards, it is extraordinary that this conversion has not been subject to serious scrutiny.[2] Why would a deviating coalition necessarily expect that the remaining set of players would act in so malevolent a fashion, without regard to their own interests?

While this point is easy enough to make and appreciate, it is somewhat less clear what one puts in its place. What one needs, in short, is a theory of intercoalitional interaction. While no particular solution is perhaps fully satisfactory, we proceed without further ado to out point of view on this matter. Imagine that, for some reason (to be endogenized later), we are faced with a *coalition structure*, a partition of the set of players into disjoint subsets. The partition means, by definition, that players within a subset are free to write arbitrary binding agreements, while players across subsets are not. In that case, we may consider the noncooperative game induced across subcoalitions, by treating each subcoalition as a player with an incomplete preference ordering. The set of all payoffs for a given coalition would then be the set of all payoffs under the Nash equilibrium of this game. See Ray and Vohra (1997) for details of this conversion, as well as Ichiishi (1981) and Zhao (1992) in a different context.

Moreover, if the underlying strategic game has interpersonally comparable utilities, and if side payments can be made across subsets of players, without affecting the strategic choices of the other players, the set of all payoffs to a coalition could then be identified with a single number, its *worth*.

In this way, we arrive at a *partition function*, one that assigns to each coalition, *and each coalition structure of which that coalition is a member*, a

[1] This particular variant is called the α-characteristic function. There are other ways to get to a characteristic function as well, with the same associated problems.

[2] Of course, the point has not passed unnoticed (see Lucas, 1963; Thrall and Lucas, 1963; Rosenthal, 1972; among others). But these papers largely restrict themselves to studying the analogs of well-known solution concepts for characteristic functions, and do not focus on endogenous coalition formation in this context.

worth, or, more generally, a set of payoffs. Given this function, we are then faced with the question: which agreements will be written and which coalition structure will form? We emphasize the *simultaneous* determination of coalition structure and payoff division among players.

1.2. *Main Features*

Partition functions permit us to get a handle on what might follow a coalitional deviation, though there are limitations.[3] To see this, consider the following examples.

EXAMPLE 1.1. Three Cournot oligopolists produce output at a fixed unit cost, c, in a homogeneous market with a linear demand curve; $p = A - bx$. They are free to form coalitions among themselves, and this includes the option of forming the grand coalition of all three players. Recall that, by standard calculations, the Nash profit accruing to a single firm in an n-player Cournot oligopoly is

$$\frac{(A-c)^2}{b(n+1)^2} = \frac{K}{(n+1)^2},$$

where $K \equiv (A-c)^2/b$. Now suppose that the three firms in our example are deciding whether or not to form a cartel. If they do, they will earn monopoly profits, which from the expression above equals $K/4$. Now it must be the case that in the proposed agreement between the three at least one of the firms is earning no more than $K/12$. What should this firm do?

The α-characteristic function tells us that if this firm breaks off, it should anticipate whatever it is that the other firms can hold it down to. But this last number is 0, for it is certainly the case that the other two firms can flood the market and drive prices down to 0. So the α-characteristic function predicts that our firm should not object to *any* nonnegative return, however small. This is clearly absurd.

On the other hand, suppose that our firm anticipates that in the event of its defection, the other two firms will play a best response to the defector's

[3] For instance, what if the game so constructed has not one Nash equilibrium but many? For more discussion, See Ray and Vohra (1997). Partition functions also fail to capture "network relationship" across players as in Jackson and Wolinsky (1996) and Dutta *et al.* (1995). Finally, there is the questions of "separability": this approach would be invalid if the outcomes leading up to the partition function were themselve conditioned in some way on the process of coalition formation. Ray and Vohra (1997) discuss this issue as well.

subsequent actions. This implies that following the deviation, we are in a duopoly, where the deviant's return, using the general expression above, is $K/9$. This exceeds $K/12$.

Does this mean a deviation from the three-player coalition is then justifiable? Not really: there are other considerations. Study the situation facing the two remaining firms once our deviant leaves. Their *total* return is $K/9$ as well, which means, of course, that one of them can be earning no more than $K/18$. If this firm were to leave and induce the standard three-person oligopoly, its return would be $K/16$. So faced with the irrevocable departure of one firm from the original agreement, the remaining firms will split up as well. But in the case, the original deviant gets $K/16$ too! So each member of the three-firm coalition would anticipate receiving $K/16$ as a result of such a deviation. It follows that the grand coalition in this example is a stable coalition structure (proposing the joint monopoly outcome with each firm getting at least $K/16$).

EXAMPLE 1.2. Consider the provision of a public good by three symmetric agents. Describe the partition function in the following intuitive way. Assume that if the three players get together, they produce a per-capita utility of 1. If one player leaves, assume that he would get 2 by free-riding on the other players' provisions, *provided* that the other two players stay together. Thus far this is analogous to the Cournot model. What is different is that we consider a case where the remaining two players will indeed wish to stay together. Imagine that by doing so, they can get a per-capita utility of 0.25. If all three players are on their own, assume that no public good is produced and that each player gets 0.

In this case, and in contrast to Example 1.1, a single deviant can credibly expect to get 2, simply because, faced with the deviation, the remaining agents will find it in their best interest to cling together. Now we have a problem, because it is clear that in the grand coalition, at least one player must get strictly less than 2. We find it difficult, in this case, to avoid an inefficient outcome.[4]

[4] This statement is fraught with numerous complexities that we have found best to avoid, in the interests of making some progress on the question of coalition formation. If any binding agreement can, in principle, be renegotiated, then the outcome should be efficient. After all, if as in the example above, one player is leaving, the other two can try to lure him back with the promise of a better offer, as the grand coalition always enjoys the advantage of superadditivity. But what gives this player a credible bargaining advantage in the first place, unless he does exercise the option to leave? And what is to guarantee that once this advantage is relinquished by his voluntary return to the grand coalition, that it will not indeed pass to someone else? It turns out that these features are not easy to model, and they possibly involve an explicit accounting for the underlying dynamics. Once these points are recognized, it becomes clear that the particular approach we follow in this paper is only one of many.

These examples illustrate two main features of our analysis. First, as discussed above, there is the question of intercoalitional *interaction* that characteristic functions neglect. This interaction is fundamental to our discussion of coalition structure.[5]

The second feature is one of consistency or "farsightedness": a player or group of players breaking off negotiations must do more than simply presume that they will be engaged in a noncooperative game with the resulting complementary coalition. They must attempt to *predict* the coalition structure that arises and not just assume that the complement will stay together. The two examples illustrate two entirely different outcomes, one in which the "short-run" belief that the complement will be unaffected is indeed vindicated, and another in which it is not.[6] As Aumann and Myerson (1988) observe,

> When a player considers forming a link with another one, he does not simply ask himself whether he may expect to be better off with this link than without it, given the previously existing structure. Rather, he looks ahead and asks himself "Suppose we form this new link, will other players be motivated to form further new links that were not worthwhile for them before? Where will it all lead? Is the *end result* good or bad for me?"

1.3. *A Summary*

Our approach to interplayer negotiation is based on Rubinstein (1982) and Chatterjee *et al.* (1989, 1993).[7]

For us, the partition function is a primitive, with the idea that underlying this function is a game in strategic form. On this partition function is defined a noncooperative bargaining game. Proposers offer to form coalitions and to divide coalitional worth in particular ways. Responders agree or disagree. Coalitions from through the course of this bargaining process.

We explicitly recognize that the problem of coalition formation is intimately linked to the problem of *which* agreements will be written among the members of the formed coalition.

[5] In this respect, we follow Bloch (1996), Chwe (1994), and Ray and Vohra (1997). For other literature on coalition structure, see for example, Shenoy (1979), Hart and Kurz (1983), Jackson and Wolinsky (1996), and Dutta (1995). But the solution concepts here do not take into account the entire chain of reactions that might follow the formation of a particular coalition. This is the second main feature of our analysis (see the main text). Ray and Vohra (1996) contains a more detailed discussion of related literature.

[6] Related "consistency" or "prediction" issues are studied in Aumann and Myerson (1988), Bloch (1996), Chakravorti and Kahn (1991), Chwe (1994), Dutta *et al.* (1989), Greenberg (1990), Ray (1989), and Ray and Vohra (1997).

[7] For related literature on bargaining, see Binmore (1985), Hart and Mas-Colell (1996), Krishna and Serrano (1996), Moldovanu (1992), Okada (1996), Perry and Reny (1994), Selten (1981), and Winter (1993).

We consider the stationary (or Markov) subgame perfect equilibrium of the bargaining game.[8] We begin by establishing an existence theorem for such equilibria (Theorem 2.1). Our theorem requires some mixing in equilibrium, but at most in the choice of coalitions that a proposer might propose to. The Appendix carefully studies the need for mixing, and shows that mixing is closely related to unacceptable proposals being made in equilibrium.

Our main results revolve around the fact that we unearth a *particular* coalition structure, with the property that such a structure is predicted by a broad class of equilibria. The analysis for general partition function games is quite complex. We therefore first present these results for games generated by *symmetric* partition functions (Section 3). The general model is then studied in Section 4. But, in principle, it is possible to read Section 4 before Section 3.

In Section 3 we begin by developing an algorithm that generates a particular coalition structure from any symmetric partition function (Section 3.1). We then argue that under *every* equilibrium in which acceptable proposals are made at each stage (with positive probability), the coalition structure given by the algorithm must result (Theorem 3.1). We provide an example in which there is an equilibrium with unacceptable proposals made, and the coalition structure of the algorithm does *not* emerge. In this sense lack of delay turns out to be fundamental to our predictions.

We provide sufficient conditions for the existence of a no-delay equilibrium (Theorem 3.2). These are *necessary* as well for the existence of a pure-strategy no-delay equilibrium (Theorem 3.3). We show by example, however, that other equilibria (with different coalition structures) might coexist. A strengthening of the existence condition gives us more: that the coalition structure predicted by our algorithm is the only one that can arise in equilibrium (Theorem 3.4). We apply these findings to the Cournot oligopoly.

Section 4 takes up the general case. Our goal here is to develop a parallel for the main result of Section 3: that no-delay equilibria predict a class of coalition structures that can be computed in a finite number of steps from the parameters of the model. While the predictions here are not as sharp as in the symmetric case (and we explain why), significant progress can be made (Theorem 4.1).

[8] The game, as described, has a plethora of subgame-perfect equilibria when there are three or more players, and this is true even for the special case of characteristic functions (Chatterjee *et al.*, 1993, Prop. 0). There is no logically convincing way to rule out such equilibria. Rather, we view Markov perfection (as many other authors do) as an interesting and perhaps focal mode of behavior, of interest in its own right.

We acknowledge the insights of Bloch (1996), which is closely related to the present exercise. His paper is motivated by very much the same questions, and explicitly studies partition functions as well. Indeed, Bloch makes use of some of the results presented in an earlier version of the paper (though his original work is quite independent of ours).

An important difference is that Bloch assumes that coalitional worths are distributed among the members according to some fixed rule. In contrast, we make no such assumption but try and deduce both coalitional structure as well as intracoalitional allocation from the *same* game. Nevertheless, in the special case of symmetric games, our results can be viewed as a vindication of Bloch's assumptions, though we obtain a somewhat sharper prediction regarding coalition structure. The general case yields additional insights.

2. A GENERAL MODEL

2.1. *The Bargaining Game and Equilibrium*

$N = \{1, \dots, n\}$ is the set of *players*. A *coalition structure of* N is a partition π of N. Let Π denote the set of all coalition structures. A *partition function* assigns to each coalition S in a coalition structures π a worth $v(S, \pi)$. Assume that $v(\{i\}, \pi) \geq 0$ for all $i \in N$ and $\pi \in \Pi$ with $\{i\} \in \pi$. Let $v \equiv \{v(S, \pi)_{S \in \pi}\}_{\pi \in \Pi}$.

In our model, players will make proposals to coalitions and respond to proposals made to coalitions to which they belong. To each coalition S is assigned an initial proposer $\rho^p(S)$, in case S is the remaining set of players in the game, and an order of respondents $\rho^r(S)$ in case $S \cap \{i\}$ is proposed to by some player i. In the latter case ρ^r is just a permutation of the players of S. The collection $\rho \equiv \{\rho^p(S), \rho^s(S)\}_{S \subseteq N}$ will be referred to as a *protocol*. A *bargaining game* is a collection $\{N, v, \rho\}$.

Interpret a bargaining game as follows. The initial proposer $\rho^p(N)$ starts the game. She chooses a coalition S (of which he is a member), and then makes a proposal to this coalition.

Loosely speaking, a proposal is the division of the worth of a coalition among its members. But given a partition function, a worth is not well defined until a coalition structure has formed in its entirety. Therefore a proposal must consist of a set of *conditional statements* that describe how the division of a coalition's worth occurs in every contingency.

The notion of a contingency here is ambiguous: it could be as minimal as the simple realization of the coalition's worth, but, in principle, it could include information such as the *process* leading up to that worth—the coalition structure formed, the order of coalition formation, and so on. In this paper we study stationary strategies, those in which active players only condition their actions on the current payoff-relevant state (a precise description will be provided below). In particular, we will adopt the narrower view of a proposal simply as a description of worth allocation for every conceivable coalition structure that finally forms. If some coalitions have already left the game, then a proposal is conditioned only on those coalition structures that are consistent with this fact.

To describe this precisely, let $\Pi(S)$ be the collection of all partitions of a coalition S. If a collection of coalitions π has left the game, then a *proposal* is a pair (S, y), where $y \equiv \{y(\pi')\}_{\pi'}$ such that $\pi' = (\pi, S, \hat{\pi})$ for $\hat{\pi} \in \Pi(N \setminus (\pi \cup S))$, and, for every such π', $y(\pi') \in \mathbb{R}^s$ is feasible in the sense that

$$\sum_{i \in S} y_i(\pi') = v(S, \pi').$$

Once a proposal (S, y) is made by a proposer i, attention shifts to the respondents in S, the order of which is obtained from $\rho^r(S)$ (with i, the proposer, eliminated from the list). By a response we mean simply an acceptance or rejection of the going proposal. If all respondents accept, the players in S retire from bargaining, and the game shifts to the set of players remaining in the game. If the set of remaining players is T, the next proposer is $\rho^p(T)$.

It remains to describe what occurs in the case of a rejection. In that case, it is assumed that the first rejector gets to make the next proposal. In addition, there is assumed to occur (as in Rubinstein, 1982) the lapse of a certain amount of time, which imposes a geometric cost on all players, and is captured by a common discount factor $\delta \in (0, 1)$. After the next proposal is made, the game continues exactly as described above. A schematic description of the extensive form is provided in Figure 1.

If and when all agreements are concluded, a coalition structure forms. Each coalition in this structure is now required to allocate its worth among its members as dictated by the proposals to which they were signatories. If bargaining continues forever, it is assumed that all players receive a payoff of 0.

A (*stationary*) *strategy* for a player requires her to make a proposal whenever it is her turn to propose, where the (possibly probabilistic) proposal is conditioned only on the current state of the game—the current player set and the coalitions that have already formed. It also requires her to accept or reject proposals at every node where she is supposed to respond. Again we impose the restriction that this (possibly probabilistic)

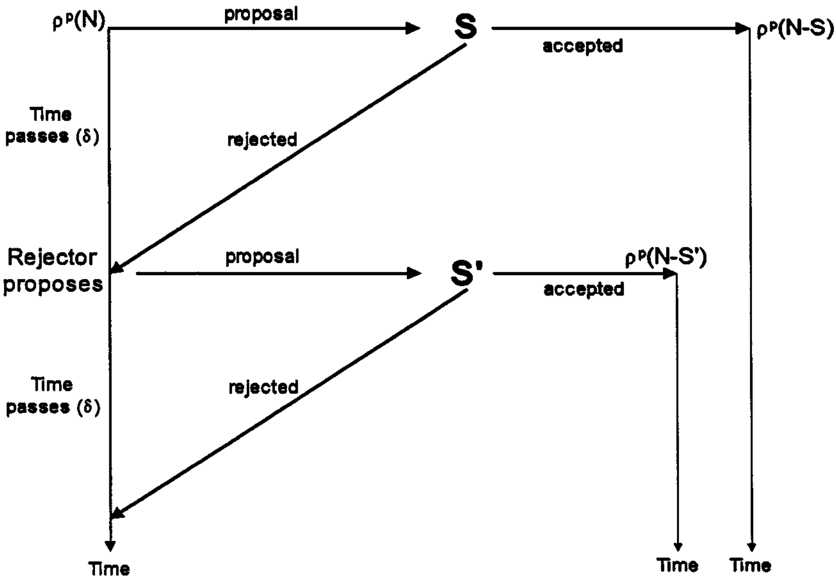

FIG. 1. A schematic description of the bargaining process.

decision not depend on anything else but the current set of players, the coalitions that have already left, as well as the identity of the proposer and the nature of the proposal that she is responding to.[9]

A *stationary (perfect) equilibrium* is defined to be a collection of stationary strategies such that there is no history at which a player benefits from a deviation from her prescribed strategy.

2.2. *Existence of Equilibrium*

Note that our notion of equilibrium allows for mixed (behavior) strategies in three ways: (a) the proposer may randomly choose a coalition, (b) given the choice of a coalition, the proposer may randomly choose offers, and (c) respondents may mix over accepting and rejecting a proposal.

But it turns our that an equilibrium exists with a minimal need to randomize, as described in the theorem below.

THEOREM 2.1. *There exists a stationary equilibrium where the only source of mixing is in the (possibly) probabilistic choice of a coalition by each proposer.*

[9] Of course, it is only fair to also let her condition her yes–no decision on the identity and order of the other respondents, but this is already accounted for, because once the proposer and proposal are given, the protocol fixes the order of respondents.

Remark 2.1. In the Appendix, we show that this theorem cannot be strengthened to assert the existence of a pure-strategy equilibrium without additional assumptions.

Remark 2.2. While the proof of this theorem (as well as the proofs of all other results) is postponed to Section 5, the argument may be of intrinsic interest. The proof relies on an inductive fixed-point argument. At every subgame, a suitable fixed point (in payoff space) is constructed, and this fixed point replaces the relevant portion of the game, as we inductively move to an earlier subgame. To complete the fixed-point argument for the earlier subgame, we need an additional continuity argument for the recursively constructed fixed points, which is where the possibility of mixing makes an appearance.

Remark 2.3. The existence argument can be readily modified to include NTU partition function games that are strictly comprehensive; see the remark following the proof.

3. SYMMETRIC PARTITION FUNCTIONS

A partition function is *symmetric* if the worth of a particular coalition in a given partition depends *only* on the number of individuals in each coalition in that partition. The vector of integers that capture this information may be referred to as a *numerical coalition structure*. More formally, let $\pi = \{S_1, \dots, S_k\}$ be a coalition structure. With some abuse of notation, the worth of a coalition $S_i \in \pi$, $v(S_i, \pi)$, can be written as $v(s_i, \mathbf{n}(\pi))$. Here $\mathbf{n}(\pi) \equiv (s_1, \dots, s_k)$, where $s_j = |S_j|$ for all j, is the numerical coalition structure associated with π.[10]

We begin the analysis by constructing a particular numerical coalition structure.

3.1. *An Algorithm*

Our results make essential use of a simple recursive algorithm which we now describe.

For a vector $\mathbf{n} = (n_i)$ of positive integers, define $K(\mathbf{n}) \equiv \sum n_i$. Use the notation ϕ to refer to the "zero-dimensional" or null vector containing no entries, and set $K(\phi) = 0$. Let $\mathcal{F}$ be the family of all such vectors (including ϕ) satisfying the additional condition that $K(\mathbf{n}) < n$.

[10] Whenever we need to emphasize the difference between a coalition and the number of players in the coalition, we will use uppercase letters to denote the coalition and lowercase letters to denote the number of players in it.

We are going to construct a rule $t(\cdot)$ that assigns to each member of this family a positive integer. By applying this rule repeatedly starting from ϕ, we will generate a particular numerical coalition structure, to be called $\mathbf{n}^*$.

Step 1. For all $\mathbf{n}$ such that $K(\mathbf{n}) = n - 1$, define $t(\mathbf{n}) \equiv 1$.

Step 2. Recursively, suppose that we have defined $t(\mathbf{n})$ for all $\mathbf{n}$ such that $K(\mathbf{n}) = m + 1, \ldots, n - 1$, for some $m \geq 0$. Suppose, moreover, that $K(\mathbf{n}) + t(\mathbf{n}) \leq n$. For any such $\mathbf{n}$, define

$$c(\mathbf{n}) \equiv (\mathbf{n} \cdot t(\mathbf{n}) \cdot t(\mathbf{n} \cdot t(\mathbf{n})) \cdots),$$

so that $K(c(\mathbf{n})) = n$, where the notation $\mathbf{n} \cdot t_1, \ldots, t_k$ simply refers to the numerical coalition structure obtained by concatenating $\mathbf{n}$ with the integers $t_1, \ldots, t_k$.

Step 3. For any $\mathbf{n}$ such that $K(\mathbf{n}) = m$, define $t(\mathbf{n})$ to be the *largest* integer in $\{1, \ldots, n - m\}$ that maximizes the expression

$$\frac{v(t, c(\mathbf{n} \cdot t))}{t}. \tag{1}$$

Step 4. Complete this recursive definition so that t is now defined on all of $\mathscr{F}$. Define a numerical coalition structure of the entire set of players N by

$$n^* \equiv c(\phi).$$

This completes the description of the algorithm.

A verbal description may be useful. Given any departed numerical coalition structure $\mathbf{n}$, which we may think of as a *substructure*, and a remaining set of players, it is possible to conceive of some *final* coalition structure that will form, for every coalition size that may be formed in this situation. (This "final" structure requires a recursive argument, as described above.)

With this scenario in mind, find those coalition sizes the *maximize the average worth of a coalition*, as described in (1). If there is more than one size, choose the *largest* coalition that achieves the desired outcome.

3.2. *Results*

3.2.1. *A Class of Equilibria That Yield n**

The departure of some given collection of coalitions induces a *stage*, defined as the set of all subgames in which a proposal is to be made, following the departure of these coalitions. For a stage in which π is the collection of coalitions that has left the game, we will denote by $\mathbf{n}(\pi)$ the numerical coalition structure corresponding to π.

For each such stage with numerical structure $\mathbf{n}$, define

$$a(\mathbf{n}) \equiv \frac{v(t(\mathbf{n}), c(\mathbf{n} \cdot t(\mathbf{n})))}{t(\mathbf{n})}. \tag{2}$$

The results of this section will depend on the following regularity condition, which will be in force throughout:

For every $\mathbf{n}$ such that $K(\mathbf{n}) < n - 1$, there exits $s \leq n - K(\mathbf{n})$ such that $v(s, \mathbf{n} \cdot s \cdot \mathbf{n}') > 0$ for all $\mathbf{n}'$ such that $K(\mathbf{n} \cdot s \cdot \mathbf{n}') = n$.

This condition implies that for all $\mathbf{n}$ such that $K(\mathbf{n}) < n - 1$, $a(\mathbf{n}) > 0$. This is the implication that needs to be kept in mind.[11]

THEOREM 3.1. *There exists $\delta^* \in (0,1)$ such that for all $\delta \in (\delta^*, 1)$, any equilibrium in which an acceptable proposal is made with positive probability at any stage must be of the following form. At a stage in which π has left the game and $\mathbf{n} = \mathbf{n}(\pi)$ belongs to $\mathscr{F}$, the next coalition that forms is of size $t(\mathbf{n})$ and the payoff to a proposer is*

$$a(\mathbf{n}, \delta) \equiv \frac{v(t(\mathbf{n}), c(\mathbf{n}))}{1 + \delta[t(\mathbf{n}) - 1]}. \tag{3}$$

In particular, the numerical coalition structure corresponding to any such equilibrium is $\mathbf{n}^*$.

Theorem 3.1 shows that if acceptable offers are made (with some positive probability) at every stage, the equilibrium coalition structure of the bargaining game *must* yield the same numerical coalition structure as our algorithm.

Thus the possibility of *delay* seems to be important in singling out the coalition structure that we identify. Delay is equivalent to the making of absurd offers which the proposer knows will be rejected. Why would such offers every be made? The answer is that a proposer may wish to pass the buck to another player, and benefit from possibly higher payoffs in some subgame. But even if this is so, can't the theorem be extended to cover such cases? To answer these questions, consider an example.

[11] If we insist on accommodating games with $a(\mathbf{n}) = 0$ for some $\mathbf{n}$, the equilibria that we identify in the main text continue to be equilibria for such games. However, the uniqueness of equilibrium cannot be expected, for obvious reasons.

The following five-player partition function will be used[12]:

$$v(4,1) = (6,2) \qquad v(3,2) = (3,8) \qquad v(2,1,1,1) = (0.1,3,3,3),$$
$$v(3,1,1) = (10,0,0) \qquad v(\pi) = 0 \qquad \text{for all other } \pi,$$

where, to ensure that the regularity condition holds, 0 should be interpreted as some small positive number.

By applying the algorithm, it is easy to check that $\mathbf{n}^* = (4,1)$.

EXAMPLE 3.1. For all discount factors sufficiently close to unity, there is an equilibrium with coalition structure $(4,1)$ in which one player makes an unacceptable proposal (in the presence of all five players) and the other four make acceptable proposals to each other. Under the equilibrium, the intransigent player receives 2δ whenever it is his turn to propose to the grand coalition. The other receive only $6/(1+3\delta)$ in their roles as proposer. We leave the details of equilibrium construction to the reader.

It follows that Theorem 3.1 can be strengthened somewhat.[13] But it cannot be strengthened free of charge:

EXAMPLE 3.2. There is also an equilibrium with coalition structure $(3,2)$. It is constructed as follows. Players 1, 2, and 3 make acceptable offers to each other and the other two make unacceptable offers to player 1. Let $\bar{x}_i$, the equilibrium payoff to i if i starts the game, be defined as

$$\bar{x}_i = \frac{3}{1+2\delta} \qquad \text{for } i = 1,2,3,$$

$$\bar{x}_j = \frac{8\delta}{1+\delta} \qquad \text{for } j = 4,5.$$

For δ close to 1, players 1, 2, and 3 get approximately 1 while players 4 and 5 get approximately 4. Clearly, player i, $i = 1, 2, 3$, cannot do better by including player 4 or 5, since $v(4,1) = (6,2)$. Given the strategies of the others, i cannot do better by making an unacceptable proposal. It is also easy to see that players 4 and 5 do not have a profitable deviation. Thus, the above strategies (together with obvious specification for nonequilibrium subgames) constitute an equilibrium.

[12] This partition function does not satisfy grand-coalition superadditivity. But it is possible to modify the example so that it does satisfy this property and so that all the examples to be based on it are valid. Details are available from the authors upon request.

[13] This possibility is taken further in Ray and Vohra (1996), where a wider class of equilibria is identified than in Theorem 3.1 here.

Thus Theorem 3.1 requires some qualification, and this qualification is associated with the possibility of delay. But this raises the following open question: is there always *some* equilibrium that yields the coalition structure $\mathbf{n}^*$?[14]

3.2.2. *No-Delay Equilibrium*

The argument above suggests that it is worthwhile to study in more detail those equilibria in which acceptable offers are made. Can we describe conditions under which they exist? Can we rule out other equilibria?

Define a *no-delay equilibrium* to be one in which, at every stage, every proposal that is made is accepted.[15]

By Theorem 3.1, we know that for discount factors close to unity, no-delay equilibria induce $\mathbf{n}^*$.

Recall that

$$a(\mathbf{n}) \equiv \frac{v(t(\mathbf{n}), c(\mathbf{n} \cdot t(\mathbf{n})))}{t(\mathbf{n})}.$$

The numbers $a(\mathbf{n})$ can, of course, be directly computed from the primitives of the model.

THEOREM 3.2. *If*

$$a(\mathbf{n}) \geq a(\mathbf{n} \cdot t(\mathbf{n})) \qquad \textit{for all } n \in F \textit{ such that } \mathbf{n} \cdot t(\mathbf{n}) \in \mathscr{F}, \tag{4}$$

then there is $\hat{\delta} \in (0, 1)$ *such that a no-delay equilibrium exists for all* $\delta \in (\hat{\delta}, 1)$.

Theorem 3.2 is useful for the following reason: if condition (4) of this theorem holds, there is always an equilibrium under which the coalition structure $\mathbf{n}^*$ identified by the algorithm must form.

Remark 3.1. It will be clear from the proof of Theorem 3.2 that when (4) holds a pure-strategy no-delay equilibrium exists. Indeed, there exists a no-delay equilibrium corresponding to every strategy in which (in every subgame following the departure of π) player i randomizes across coalitions of size $t(\mathbf{n}(\pi))$, making an acceptable proposal in each case. However, the numerical coalition structure corresponding to any no-delay equilibrium is $\mathbf{n}^*$.

[14] Ray and Vohra (1996, p. 15) discuss this issue in more detail.

[15] This is stronger than the class of equilibria identified in Theorem 3.1, but as we are after existence and uniqueness here, our results will apply *a fortiori* to the broader class.

As Theorem 3.3 indicates, condition (4) is fairly tight. It is *necessary* for the existence of a no-delay equilibrium in *pure strategies*.

THEOREM 3.3. *If there is* $\hat{\delta} \in (0,1)$ *such that a pure-strategy no-delay equilibrium exists for all* $\delta \in (\hat{\delta}, 1)$, *then* (4) *holds*.

To see how condition (4) works, consider

EXAMPLE 3.3. In the five-player partition function introduced earlier, recall that $\mathbf{n}^* = (4,1)$. So $a(\phi) = 1.5 < 2 = a(t(\phi))$, so that (4) fails. By Theorem 3.3, there is no pure-strategy no-delay equilibrium for discount factors close to unity. Indeed, no-delay equilibria fail to exist as well. To prove this, suppose, on the contrary, that such an equilibrium exists along a sequence of discount factors tending to unity. Then, by Theorem 3.1, a proposer receives $6/(1 + 3\delta)$, and makes an offer to some four-player coalition. A responder receives $6\delta/(1 + 3\delta)$.

Fix any $\delta \geq \delta^*$ such that $\delta[0.5 + 6\delta] > 6$. Now observe that there is *some* pair of individuals i and j such that if it is j's turn to propose, an offer is made to i with probability *no more* than 3/4. If individual i deviates by making an unacceptable offer to j, then the present value of i's payoff is bounded below by $\delta[(3/4)6\delta/(1 + 3\delta) + (1/4)2]$, while by sticking to equilibrium policy, he obtains $6/(1 + 3\delta)$. Comparing these two expressions under the given restriction on δ, it can easily be checked that a deviation is profitable. This completes the argument.

At the same time, (4) does not exclude the possibility that there may be other equilibria yielding entirely different coalition structures. To see this, consider

EXAMPLE 3.4. Modify the five-player partition function so that $v(4,1) = (6,1)$. Again, $t(\phi) = 4$. But now $a(4) = 1 < a(\phi)$ and it is easy to see that condition (4) holds. So there exists a no-delay equilibrium with the coalition structure (4, 1). However, the nonsymmetric equilibrium of Example 3.2, with the coalition structure (3, 2), is an equilibrium here as well.

Example 3.4 makes it clear that uniqueness needs more than condition (4). In view of Remark 3.1 it is also clear that no such strengthening of (4) can rule out mixed strategy no-delay equilibria. To state this additional requirement, we extend the definition of $t(\mathbf{n})$.

For each $n \in \mathscr{F}$ and each positive integer $l \in \{1, \ldots, n - K(\mathbf{n})\}$, define

$$\tau_l(\mathbf{n}) \equiv \arg \max_{t \in \{1,\ldots,l\}} \frac{v(t, c(\mathbf{n} \cdot t))}{t}. \tag{5}$$

In words, $t \in \tau_l(\mathbf{n})$ solves the same maximization problem as described in the algorithm, except that maximum size is restricted by l. Because of possible nonconvexities, this maximum restriction can be binding even if $l \notin \tau_l(\mathbf{n})$.[16] The utility of introducing this construction is brought out in

THEOREM 3.4. *Suppose that for each* $\mathbf{n} \in \mathscr{F}$ *and each positive integer* $l \in \{1, \ldots, n - K(\mathbf{n})\}$,

$$a(\mathbf{n}) \geq a(\mathbf{n} \cdot t) \qquad \text{for all } t \in \tau_l(\mathbf{n}). \tag{6}$$

Then there is $\hat{\delta} \in (0,1)$ *such that for all* $\delta \in (\hat{\delta}, 1)$, *every equilibrium must be no-delay, and therefore generate the numerical coalition structure* $\mathbf{n}^*$.

Okada (1996) shows that in superadditive TU games, equilibria with delay can be ruled out by modifying the bargaining game such that the proposer is chosen at random. In the context of a partition function game, superadditivity is a very restrictive assumption. However, it is easy to show that it implies uniqueness of no-delay equilibria in the present context, even without requiring proposers to be chosen at random.

A game is said to be *superadditive with respect to c* if, for any coalition structure $\pi = (t_1, \ldots, t_k)$,

$$\begin{aligned} &v\big(t_i + t_j, c(t_1 \cdots t_{i-1} \cdot t_i + t_j)\big) \\ &\qquad \geq v(t_i, \pi) + v_i(t_j, \pi) \qquad \text{for all } i, j \in \{1, \ldots, k\}, i < j. \end{aligned}$$

THEOREM 3.5. *Suppose a game satisfies superadditivity with respect to c. Then there exists* $\hat{\delta} \in (0,1)$ *such that for all* $\delta \in (\hat{\delta}, 1)$, *every equilibrium is a no-delay equilibrium, and therefore generates the numerical coalition structure* $\mathbf{n}^*$.

We end this section by observing that the conditions outlined in this section can be check in models of economic interest. An example based on the Cournot model satisfies (6) (see below), and so is the public goods model studied in Ray and Vohra (1996).

3.3. *A Cornot Oligopoly*

We apply our results to an example of a symmetric Cournot oligopoly. Suppose that n oligopolist produce a quantity x of a homogeneous product, the price P of which is determined by a linear demand curve: $P = A - bx$. Assume that there is a fixed unit cost of production, given by c.

Normalize the parameters to that $(A - c)^2/b = 1$. Using the formula for Cournot–Nash equilibrium (already presented in Example 1.1), we may

[16] Of course, $t(\mathbf{n}) \in \tau_l(\mathbf{n})$ whenever $l \geq t(\mathbf{n})$.

construct a partition function for this symmetric game. Suppose that a numerical coalition structure n forms. Consider a coalition structure of size s in this structure. Then, using our normalization and denoting by q the number of coalitions in $\mathbf{n}$,

$$v(s, \mathbf{n}) = \frac{1}{(q+1)^2}.$$

The Cournot example is quite telling in one respect. Notice how the partition function is independent of the coalition concerned, but depends *entirely* on the overall coalition structure. This feature highlights how partition functions might radically differ from characteristic functions, where all the interesting action comes from variation in coalitional worth.

The calculations in this example will draw heavily on Bloch (1996). *Assuming* equal division for coalitional worth, Bloch constructed an equilibrium coalition structure for this model. In doing so he used an algorithm similar to ours,[17] by applying Ray and Vohra (1996, Theorem 6.3) (currently condition (4) of Theorem 3.2). As we shall see, much more can be said about equilibria in this model by appealing to our results on symmetric games. We will show that this model also satisfies condition (6) and, therefore, our algorithm yields the *only* possible equilibrium coalition structure. Moreover, (approximately) equal division for high discount factors is a result rather than an assumption.

THEOREM 3.6 (Generalization of Bloch (1996)). *All equilibria in a Cournot oligopoly with n firms are no-delay equilibria. Moreover, there is a unique numerical equilibrium coalition structure. It consists of L singleton firms and a single cartel of size $n - L$, where L is the smallest nonnegative integer such that*

$$n - L < (L+2)^2 + 1.$$

Thus our results predict full cartelization in this example whenever there are four firms or less, and imperfect cartelization thereafter.

This observation can be quickly established using the algorithm of Section 3.1, and then checking that the uniqueness condition (6) of Theorem 3.4 is indeed satisfied. While the reader should consult the proof for details, it is easy to provide some intuition. To do so, we invoke an important observation due originally to Salant *et al.* (1983): if several firms are already out of a potential cartel, and the number of firms left is "small enough," then the remaining firms will not find it advantageous to form a cartel. Intuitively, the gain in market concentration does not justify the

[17] In some cases, he obtains two equilibrium coalition structures, whereas our algorithm yields a unique numerical coalition structure.

profit sharing that will be needed. Applying this idea recursively to the remaining number of players, we can find a threshold at which the average payoff to the remaining players, if they stay together, is approximately the same as when a player quits, sparking off a cartel collapse.

Summarizing so far, we see that at this threshold, firms would rather stay together than break up. But *knowing this is so*, those firms in excess of this threshold will disagree to form a cartel as well, predicting correctly that the remaining firms will stay together. This creates an equilibrium outcome with one large cartel and several singleton firms.

4. THE GENERAL CASE

The focus of the analysis for symmetric games is the identification of a particular (numerical) coalition structure, $\mathbf{n}^*$, which is generated by a "broad" class of equilibria. In particular, we showed that if an equilibrium satisfies the no-delay requirement, then it must generate $\mathbf{n}^*$ as the equilibrium coalition structure.

It is natural to ask if a corresponding observation applies in the general case. That is, can we identify a particular coalition structure, or a class of structures, such that a no-delay equilibrium will generate a coalition structure within this class? This is the task to which we set ourselves in the current section.[18]

We reiterate what we mean by the "identification of a particular [class of] coalition structure[s]." It must be possible to take the parameters of the model, and compute, in a *finite* number of steps, the relevant structure(s). The work then lies in proving that the structures are the outcome of certain equilibria.

In attempting such a generalization, three points must be noted at the very outset, and each of these stands in sharp contrast to the symmetric case. First, there is no hope of finding, in general, a *single* (numerical) coalition structure as the predicted outcome. Second, the predicted structure may well depend on the bargaining protocol. Third, the assumption of equal division (as in Bloch, 1996) may be unacceptably restrictive in general models. All these points may be illustrated by means of a single example.

EXAMPLE 4.1. We use the special case of characteristic function. $N = \{1, 2, 3\}$. Worths are given as follows: $v(\{12\}) = 3$, $v(\{123\}) = 4$, while $v(S) = 0$ for all other S. Direct computation easily verifies that there is a

[18] It should be pointed out that the analysis involves a number of subtle details, and the results are not as clear-cut. It may be worth skipping this section at a first reading, and absorbing the proofs for the symmetric case instead.

unique stationary equilibrium, and it involves no delay. For the discount factor close enough to unity, bargaining game started by player 1 or 2 will result in the formation of the coalition structure $(\{12\}, \{3\})$, where players 1 and 2 "produce" and player 3 is left out. On the other hand, if player 3 begins the game, the single coalition $\{123\}$ will form, with players dividing this worth unequally within this coalition (even as $\delta \to 1$). This observation makes three points: (1) there may be more than one equilibrium coalition structure; (2) the structure that arises may well depend on the protocol; and (3) the assumption of equal division within formed coalitions may be seriously restrictive in nonsymmetric cases. These new features are incorporated in the analysis that follows.

We begin, then by describing an algorithm that generates a class of coalition structures. The main theorem then ties no-delay equilibrium to the generation of these structures. Several steps are involved in the description of the class.

4.1. *A Class of Coalition Structures*

4.1.1. *Some Observations on Characteristic Functions*

It will be convenient to begin with some observations for characteristic functions. The analysis in this subsection extends a construction introducing in Chatterjee *et al.* (1989).

Let S be a finite set of players. A (TU) *characteristic function* w assigns a number $w(T)$ (normalized to be nonnegative) to every coalition T of S.

We continued to use lowercase letters $s, t, t_k, \ldots$ to denote the cardinalities of coalitions $S, T, T_k, \ldots$.

Our task in this subsection is to allocate, to each player $i \in S$, a number $a_i(w)$, as well as a set of coalitions, $\mathscr{C}_i(w)$.[19]

Step 1. Consider the problem

$$A^1 \equiv \max_{T \subseteq S} \frac{w(T)}{t}. \tag{7}$$

For each coalition T that solves (7), let

$$\Delta(T) \equiv -A^1 \frac{t-1}{t},$$

[19] It is useful to note that analogous construction in the symmetric case, which was embodied in the algorithm in Section 3.1. There each player was assigned the same number, as well as the same set of coalitions (those of maximal size among those maximizing "average worth").

and define $\mathscr{S}_1$ to be the collection of all T that minimize $\Delta(T)$, subject to the constraint that they solve (7). If $A^1 > 0$, clearly, this means: include T in $\mathscr{S}_1$ if and only if it solves (7) and there is not other T' that solves (7) and is of larger *size* (there is no need for T' to be a superset of T).

Let Δ^1 be the value of $\Delta(T)$ is this class

Define U_1 to be the union of all players who belong to coalitions that belong to S_1: i.e., $U_1 \equiv \{i \in S | i \in T \text{ for some } T \in \mathscr{S}_1\}$. Define

$$a_i(w) \equiv A^1 \qquad \text{for all } i \in U_1, \tag{8}$$

and a set of coalitions, for each $i \in U_1$, by

$$\mathscr{C}_i(w) \equiv \{T \in \mathscr{S}_1 | i \in T\}. \tag{9}$$

If $U_1 = S$, end here. Otherwise go on to Step 2.

Step 2. Recursively, suppose that the values $(A^1, \dots, A^k; \Delta^1, \dots, \Delta^k)$ and the coalitions $(U_1, \dots, U_k)$ have been defined for some integer $k \geq 1$, and that $S \setminus \bigcup_{j=1}^{k} U_j \equiv S' \neq \varnothing$.

Consider the problem

$$A^{k+1} \equiv \max_{T \equiv T_1 \cup \cdots \cup T_{k+1}} \frac{w(T) - \sum_{j=1}^{k} A^j t_j}{t_{k+1}}, \tag{10}$$

where the maximization takes place over coalitions of the form $T_1 \cup \cdots \cup T_{k+1}$, subject to the constraint that $T_j \subseteq U_j$ for all $j = 1, \dots, k$, and $\varnothing \neq T_{k+1} \subseteq S'$.

For each coalition of the form $T = T_1 \cup \cdots \cup T_{k+1}$ that solves (10), let

$$\Delta(T) \equiv -\frac{\sum_{j=1}^{k} t_j(\Delta^j + A^j)}{t_{k+1}} - A^{k+1} \frac{t_{k+1} - 1}{t_{k+1}}.$$

Define $\mathscr{S}_{k+1}$ to be the collection of all coalitions that minimize $\Delta(T)$, subject to the constraint that they solve (10). As in Step 1, this implies a selection (based on size) from the set of solutions to (10), but has no direct connection with maximal coalitions, as in that step.

Let Δ^{k+1} be the value of $\Delta(T)$ in this class.

Define

$$U_{k+1} \equiv \{i \in S' | i \in T \text{ for some } T \in \mathscr{S}_{k+1}\}.$$

Let

$$a_i(w) \equiv A^{k+1} \qquad \text{for all } i \in U_{k+1}, \tag{11}$$

and define a set of coalitions, for each $i \in U_{k+1}$, by

$$\mathscr{C}_i(w) \equiv \{T \in \mathscr{S}_{k+1} | i \in T\}. \tag{12}$$

If $S \setminus \bigcup_{j=1}^{k} U_j = \emptyset$, end the recursion here. Otherwise repeat Step 2.

As already indicated, this construction assigns to each individual i in the player set S, a number $a_i(w)$ as well as a set of coalitions $\mathscr{C}_i(w)$, depending on the characteristic function w defined on S.

4.1.2. *Rules of Coalition Formation*

Let N be the player set of the original game. Denote by $\Pi°$ the collection of all coalition substructures of N; i.e., the collection of all coalition structures of every strict subset of N.[20] For each $\pi \in \Pi°$, let $\mathscr{T}(\pi)$ be the collection of all coalitions formed from the remaining set of players (not included in π).

A *rule of coalition formation* (RCF) is a map R: $\Pi° \to \mathscr{T}(\pi)$. In words, given any substructure, R assigns a fresh coalition from the set of players not in the substructure.

Given any RCF, a substructure can be "completed" into a full coalition structure of N in the obvious way, by recursively applying the RCF starting from that substructure until no players are left. This induces a *completion map* from $\Pi°$ to the set Π of (full) coalition structures of N. Call this map $c(\cdot, R)$; it depends on the RCF R. It will be notationally used to define $c(\pi, R) \equiv \pi$ for all (full) coalition structures π.

4.1.3. *Characteristic Functions Induced by an RCF*

Consider a substructure π of $\Pi°$ with the property that there is a nonempty set of players $S(\pi)$ not included in π. Given some RCF R, a characteristic function $\omega_{R\pi}$ is induced on $S(\pi)$ in the following way:

$$w_{R\pi}(T) \equiv v(T; c(\pi \cdot T, R)) \tag{13}$$

for all nonempty $T \subseteq S(\pi)$.

4.1.4. *Consistent Rules of Coalition Formation*

Recall that for every characteristic function w defined on some set of players S, we have assigned a number of $a_i(w)$ and a set of coalitions $\mathscr{C}_i(w)$ to every player $i \in S$.

[20] As in the symmetric case, the "empty structure" ϕ is also an element of $\Pi°$.

Say that a rule of coalition formation R is *consistent* if, for every substructure $\pi \in \Pi^\circ$,

$$R(\pi) \in \mathscr{C}_j(w_{R\pi}), \tag{14}$$

where j is the first proposer assigned by the bargaining protocol when the set of active players is $S(\pi)$.

By simply working backwards from substructures π such that $S(\pi)$ is a singleton, it is elementary to check that a consistent RCF always exists.

It is important to note that our description of a consistent rule of coalition formation depends only on the parameters of the model.[21] Moreover, the description is finite, in the sense that, given any partition function and a bargaining protocol, every consistent RCF can be identified by using a bounded sequence of computations.

4.2. *No-Delay Equilibrium and Consistent RCFs*

Just as in the special case of symmetric games, a no-delay equilibrium is one in which, at every stage, every proposal that is made is accepted.

THEOREM 4.1. *There exists $\delta^* \in (0,1)$ such that if $\delta \in (\delta^*, 1)$, every no-delay equilibrium must generate a coalition structure given by some consistent rule of coalition formation. Formally, given a no-delay equilibrium, there exists a consistent RCF R such that every stage indexed by $\pi \in \Pi^\circ$, with the proposer given by the bargaining protocol, the coalition that is formed corresponds to $R(\pi)$. In particular, the coalition structure that emerges in equilibrium is given by $c(\phi, R)$, where ϕ corresponds to the null substructure.*

Theorem 4.1 generalizes the corresponding results obtained for symmetric games. To see this, it is sufficient to note that in symmetric games satisfying the assumption that $a(\mathbf{n}) > 0$, *every* consistent RCF yields the same numerical coalition structure, and that structure is precisely $\mathbf{n}^*$. We leave the details of the argument to the reader.

In general, of course, there may well be several coalition structures generated by the class of consistent RCFs. This is certainly true if we alter bargaining protocols (Example 4.1), and may even be true for a given protocol.

The careful reader will have noted that in the main building block of our algorithm (Section 4.1.1 on characteristic functions), not only is a maximization problem solved (the problem described in (10)), but a further refinement of the set of maximizing coalitions is needed (this is the additional selection involved in minimizing $\Delta(T)$; see the discussion follow-

[21] Unlike the case of symmetric games, we are forced here to include dependence on the bargaining protocol as well.

ing (10)). In symmetric games, there is not direct analog to this (except requiring that the coalition maximizing "average worth" be as large as possible). The following example is designed to explain this additional restriction, as well as to point out that the choice of "largest" coalitions does not carry through to the general case. Again, the example only needs a characteristic function in order to make the point.

EXAMPLE 4.2. Consider the following four-person characteristic function: $N = \{1,2,3,4\}$, $v(\{1\}) = 1$, $v(\{12\}) = 2$, $v(\{123\}) = 2.8$, $v(\{1234\}) = 3.6$, $v(S) = 0$ for all other coalitions S. Let us compute $\{a_i\}$ and $\{\mathscr{C}_i\}$ for this game. Because no partition function is involved, this is simple.

Begin with Step 1 in Section 4.1.1. We see that $A^1 = 1$. Two coalitions —$\{1\}$ and $\{12\}$—achieve this outcome. The refinement following (7) dictates that the smaller coalition be discarded. This is reminiscent of symmetric games. We thus see that $U_1 = \{12\}$. For future use, note that $\Delta^1 = -1/2$.

Now we compute U_2. By carrying out the maximization problem in (10), with (A^1, U_1) given, we see that the maximizing coalitions are $T \equiv \{123\}$ and $T' \equiv \{1234\}$. Moreover, $A^2 = 0.8$. Now observe that

$$\Delta(T) = -\frac{2(1-0.5)}{1} - 0.8\frac{0}{1} = -1,$$

while

$$\Delta(T') = \frac{2(1-0.5)}{2} - 0.8\frac{1}{2} = -0.9.$$

Our criterion of minimizing Δ therefore requires us to discard the *larger* coalition $\{1234\}$ in favor of the smaller coalition $\{123\}$. Thus $U_2 = \{3\}$. This leaves the singleton, $\{4\}$, which must obviously solve the remaining problem (nothing of interest to report here).

Thus the algorithm predicts that the coalition $\{123\}$, and not the grand coalition, will form if player 3 is to start the game, despite the fact that the grand coalition also solves the relevant maximization problem.

Direct computation of the stationary equilibrium verifies that this is indeed the case. It is easy enough to see that for δ close enough to unity, $a_1(\delta) = a_2(\delta) = 2/(1+\delta)$, $a_3(\delta) = 2.8 - 4\delta/(1+\delta)$, and $a_4(\delta) = 3.6 - 2.8\delta - 4\delta(1-\delta)/(1+\delta)$ comprise the unique solution to (46) below. It can also be checked that $\mathscr{C}_3(\delta) = \{T\}$ for δ sufficiently close to unity.

It might be argued that for "generic" games the maximization exercise described in Section 4.1.1 will have unique solutions, and therefore the additional restrictions will usually not be needed. If by genericity we mean a random (nonatomic) draw from the space of partition functions, this is certainly the case: we do have a generically unique prediction of coalition structure for any given bargaining protocol. At the same time, we hesitate to impose such genericity concepts. For example, symmetric games form an important special case, in our opinion. Yet are they "generic"?

Indeed, with issues of "genericity" neglected, it should be pointed out that the description achieved in Theorem 4.1 is not strong enough (in general). More delicate characterizations can be used to refine the predicted set even further. We omit the details.

5. PROOFS

Proof of Theorem 2.1. It will be useful to develop some additional notation. Use of notation $-S$ or $-\pi$ to denote the set of players that are left in N after the players in S or π have left. When a subcoalition S leaves the game, this defines a new bargaining game $(-S, \bar{v}, \bar{\rho})$, where $\bar{v} = \{v(T,(S,\pi))_{T \in \pi}\}_{\pi \in \Pi(N \setminus S)}\}$ and $\bar{\rho}$ is the restriction of ρ to $N \setminus S$. We will denote such games simply as $(-S, v, \rho)$. In a similar manner we can also define a game $(-\pi, v, \rho)$ corresponding to a situation in which coalitions π have left the game.

Recall that we assume that $v(\{i\}, \pi) \geq 0$ for all $i \in N$ and $\pi \in \Pi$ such that $\{i\} \in \pi$. This implies that the equilibrium payoff (if there is one) to every player is bounded above by a nonnegative number $m = \max(v(S, \pi)_{s \in \pi})_{\pi \in \Pi}$. In our search for equilibrium payoffs, we may, therefore, restrict the feasible payoff profiles to lie in X, the cube in $\mathbb{R}_+^N$ with vertex 0 and length m.

The proof is by induction on the number of players. Suppose an equilibrium exists for every game with less than n players. For the one-player model, this assumption is trivially satisfied.

In particular, the hypothesis implies that an equilibrium exists for every subgame $(-S, v, \rho)$ for every nonempty coalition S. For each such subgame fix one equilibrium strategy profile for the players of that subgame. Our goal is to describe equilibrium strategies for all the remaining nodes in the larger game that will be grafted onto the fixed strategies for the subgames.

We begin by invoking the assumption that the protocol assigns a unique continuation to the game after S has formed, regardless of how S came to be. By using the given equilibrium strategies after S forms, we may generate two objects: (i) a probability distribution β^S over $\Pi(-S)$, and (ii) a vector of expected equilibrium payoffs for all the players in $-S$, to be denoted by $u_j(S)$, for $j \in -S$.

Now consider the overall game. Let $\mathcal{N}_i$ be the set of all nonempty coalitions containing player i and let Δ_i denote the set of probability distributions over $A_i = (\mathscr{N}_i, (\{j\})_{j \in N \setminus \{i\}})$. Recall that i can only make proposals to coalitions in $\mathscr{N}_i$. It can also make an unacceptable proposal to player j (there is no loss of generality in assuming that it cannot make unacceptable proposals to other coalitions). Now α_i will denote player i's choice concerning coalitions to form or other players to whom an unacceptable offer is made. More precisely, we will interpret $\alpha_i(S)$ to be the probability with which i chooses to make an acceptable proposal to $S \in \mathscr{N}_i$ and $\alpha_i(\{j\})$ to be the probability with which i chooses to make an unacceptable proposal to player j.

Define $\Delta \equiv \prod_{i \in N} \Delta_i$. Fix a vector $\alpha \in \Delta$, and a vector $x \in X$, the latter to be interpreted below as the vector of expected equilibrium payoffs that each player receives in the game, if i is the first proposer. Consider player i. The following options are available.

First, i can name a coalition S in $\mathscr{N}_i$, and make a proposal $y(S, \pi)$ conditioned on each $\pi \in \Pi(-S)$. This will be interpreted in the sequel as an acceptable proposal. Consider the problem:

$$\max_y \sum_{\pi \in \Pi(-S)} \beta^S(\pi) y_i(S, \pi) \tag{15}$$

subject to the constraints

$$\sum_{\pi \in \Pi(-S)} \beta^S(\pi) y_j(S, \pi) \geq \delta x_j \qquad \text{for all } j \in S, j \neq i, \tag{16}$$

$$\sum_{j \in S} y_j(S, \pi) \leq v(S, (S, \pi)) \qquad \text{for each } \pi \in \Pi(-S). \tag{17}$$

Denote by $g_i(S, x)$ the maximum value so attained. It is easy to see that

$$g_i(S, x) = \sum_{\pi \in \Pi(-S)} \beta^S(\pi) v(S, (S, \pi)) - \delta \sum_{j \in S; j \neq i} x_j,$$

which is clearly a continuous function of x, and is independent of α.

Second, i might make an unacceptable proposal to j.

Both these cases can be considered together in the following way. What we will do is compute a particular present value payoff to i, in a situation where $(x, \alpha) \in X \times \Delta$ is given. We will show thereafter that i's attempt to maximize this value, with respect to his choice of proposal probabilities, yields an equilibrium response. For a *fixed* i, define a collection $\{v_i^j(x, \alpha)\}_{j \in N}$ in the following way:

$$v_i^j(x, \alpha) = B_i^j + \delta \sum_{k \neq j} \alpha_j(\{k\}) v_i^k(x, \alpha)$$

for all j and k, where

$$B_i^i \equiv \sum_{S \in \mathscr{N}_i} \alpha_i(S) g_i(S, x),$$

and, for $j \neq i$,

$$B_i^j \equiv \delta x_i \left[\sum_{S \in \mathscr{N}_j;\, i \in S} \alpha_j(S) \right] + \sum_{S \in \mathscr{N}_j;\, i \notin S} \alpha_j(S) u_i(S).$$

We may interpret v_i^j as the *value* that player i receives when player j proposes at this stage. Note, for future use, that the value is taken to depend on i's best payoff $g_i(S, x)$ from making an acceptable proposal to each coalition S, as well as the entire vector α.

The set of simultaneous equations defining the vector $V_i = (v_i^j)$ can be written in matrix form as $CV_i = B_i$, where $B_i = (B_i^j)$ and C is the $n \times n$ matrix with 1's on the diagonal and $-\delta\alpha_j\{k\}$ as the jkth off-diagonal element. Note that the sum of the off-diagonal elements in any row lies in the half open interval $(-1, 0]$ and C is the nonsingular. It is now easy to see that v_i^j is continuous in x and α for all j.

Now define a function on $X \times \Delta \times \Delta_i$ by

$$v_i(x, \alpha, \alpha_i') \equiv \sum_{S \in \mathscr{N}_i} \alpha_i'(S) g_i(S, x) + \delta \sum_{j \neq i} \alpha_i'(\{j\}) v_i^j(x, \alpha), \quad (18)$$

and maximize this function with respect to $\alpha_i' \in \Delta_i$.

Let $\phi_i^1(x, \alpha)$ denote the maximum value of this problem, and let $\phi_i^2(x, \alpha)$ denote the set of maximizers. It is easy to see, using the maximum theorem and the fact that $v_i(x, \alpha, \alpha_i')$ is continuous, that $\phi_i^1(x, \alpha)$ is a continuous function and that $\phi_i^2(x, \alpha)$ is a convex-valued, upper hemicontinuous correspondence. Since $v(\{i\}, \pi) \geq 0$ for all i and $\pi \in \Pi$, it follows that, for all $(x, \alpha) \in X \times \Delta$, $\phi_i^1(x, \alpha) \in [0, m]$ for all i. Thus $\prod_i \phi_i^1$ maps from $X \times \Delta$ into X. Therefore the correspondence

$$\phi \equiv \prod_i \phi_i^1 \times \prod_i \phi_i^2 \colon X \times \Delta \mapsto X \times \Delta$$

satisfies all the conditions of Kakutani's fixed-point theorem and has a fixed point $(\bar{x}, \bar{\alpha})$.

We shall now use this fixed point to construct an equilibrium. Let σ denote the strategy profile such that:

(i) When the player set is N, player i as a proposer makes proposals according to $\bar{\alpha}_i$. To every coalition $S \in \mathscr{N}_i$ such that $\bar{\alpha}_i(S) > 0$, she proposes $y(S, \pi)$ which solves the problem defined by (15), (16), and (17).

To every $j \neq i$ such that $\overline{\alpha}_i(\{j\}) > 0$, she offers, for every possible partition containing the coalition $\{i, j\}$, less than $\delta \bar{x}_j$ (possibly negative). This yields player i a payoff of $\bar{x}_i$.

(ii) Suppose the player set is N, player i is a respondent to a proposal $y(S, \pi)$, and every respondent j to follow i is offered an expected payoff at least δx_j, i.e., $\sum_{\pi \in \Pi(-S)} \beta^S(\pi) y_j(S, \pi) \geq \delta \bar{x}_j$ for all respondents j that follow i. Then i accepts the proposal if and only if

$$\sum_{\pi \in \Pi(-S)} \beta^S(\pi) y_i(S, \pi) \geq \delta \bar{x}_i.$$

(iii) Suppose the player set is N, and player i is the respondent. From (ii) we know that if there is exactly one respondent to follow i, say player j, such that j is offered an expected value less than $\delta \bar{x}_j$, then j will reject the proposal. Player i's decision will now depend on the present value of the payoff to i resulting from j rejecting the offer and making a proposal as in (i). In fact, this values is precisely $\delta v_i^j(\bar{x}, \overline{\alpha})$. Player i accepts the proposal if and only if

$$\delta v_i^j(\bar{x}, \overline{\alpha}) \geq \delta \bar{x}_i.$$

Note that this inequality might hold even though we know from the construction of v_i^i and the fact that $(\bar{x}, \overline{\alpha})$ is a fixed point, that $\delta v_i^j(\bar{x}, \overline{\alpha}) \leq \bar{x}_i$. Now consider a proposal made to respondents $\{1, \ldots, r\}$, *in the given order*. Inductively, suppose we have computed in the decisions of all respondents $i + 1, \ldots, r$. Player i's decision is then obtained by considering the decision of the next responder, say j, who rejects the proposal. Player i accepts the proposal if and only if $\delta v_i^j(\bar{x}, \overline{\alpha}) \geq \delta \bar{x}_i$. In this way we obtain a complete description of the actions of all respondents of a proposal.

(iv) If the player set is not N, it must result from some collection of coalitions π having left the game. The strategies of the remaining players are defined according to the preselected equilibrium of the game $(-\pi, v)$.

We can now show that a strategy profile σ satisfying (i)–(iv) is a stationary equilibrium. Consider such a strategy and deviations that a single player i can contemplate. By construction, $\bar{x}_i = v_i(\bar{x}, \overline{\alpha}, \overline{\alpha}_i) = \max v_i(\bar{x}, \overline{\alpha}, \cdot)$. This means that it is not possible for i as a proposer to receive a higher payoff than $\bar{x}_i$ by making a one-shot deviation from $\overline{\alpha}_i$. This implies that no other strategy can yield i a higher payoff than $\bar{x}_{\mathrm{i}}$. The action prescribed in (i) achieves $\bar{x}_i$ and, therefore, cannot be improved upon. Suppose i is a respondent and all respondents to follow i are offered at least $\delta \bar{x}_j$, which, by hypothesis, they will accept. By rejecting the proposal i gets a present value of $\delta \bar{x}_i$. Clearly, then, the action prescribed in (ii) cannot be improved upon. Suppose i is a respondent who is followed by a respondent j who, based on σ, will reject the proposal. Accepting the proposal yields $\delta v_i^j(\bar{x}, \bar{\alpha})$ to player i while rejecting it yields at most $\delta \bar{x}_i$. Thus the action described in (iii) cannot be improved upon. A similar argument applies to the description in (iii) of i's actions in the other cases when i is a responder. Finally, note that when some players have left the game, the actions in (iv) are obviously unimprovable. Thus, σ is a stationary equilibrium. ■

Remark. The existence argument can be modified to include NTU partition function games that are strictly comprehensive. Let $V(S, \pi) \subseteq \mathbb{R}^S$ denote the utility set of a coalition S, under the coalition structure π. Condition (17) in the proof of Theorem 2.1 will now have to be changed to

$$y(S, \pi) \in V(S, (S, \pi)) \qquad \text{for each } \pi \in \Pi(-S).$$

If $V(S)$ is closed and strictly comprehensive, it is easy to check that $g_i(S, \cdot)$ is well defined and, by the maximum theorem, a continuous function. (Of course, it can no longer be written in the simpler form used in the proof of Theorem 2.1.) The rest of the proof remains unchanged. It is not possible to weaken strict comprehensiveness to comprehensiveness. If the utility sets are not strictly comprehensive, it is possible that conditions (16) and an appropriately modified version of (17) cannot simultaneously be satisfied; i.e., $g_i(S, x)$ need not be well defined. Moreover, even if the problem of maximizing the expression in (18) is defined only to cover those cases in which $g_i(S, x)$ is well defined, the maximum value need not be continuous. In fact, an equilibrium may not exist; see Example 2.6 of Bloch (1996).[22]

To prove Theorem 3.1, our first task is to fix δ^*. This is done by the help of the following result.

LEMMA 5.1. *There exists $\delta^* \in (0, 1)$ such that for any $\delta \in (\delta^*, 1)$ and any $\mathbf{n} \in \mathscr{F}$, $t(\mathbf{n})$ is the unique integer in the set $\{1, \dots, n - K(\mathbf{n})\}$ that maximizes*

$$\frac{v(t, c(\mathbf{n} \cdot t))}{1 + \delta[t - 1]}. \tag{19}$$

Proof. For $\mathbf{n}$ such that $K(\mathbf{n}) = n - 1$, the statement is trivially true. Fix, therefore, some $\mathbf{n} \in \mathscr{F}$ such that $K(\mathbf{n}) < n - 1$ and consider a sequence $\{\delta^q\}$ in $(0, 1)$ such that $\delta^q \to 1$. Let $\mu(\mathbf{n}, \delta^q)$ denote the set of

[22] While Bloch considers only pure-strategy equilibria, it can be shown that in this example there do not exist any mixed-strategy equilibria either.

maximizers (in t) of the expression in (19) corresponding to δ^q. By the maximum theorem, this correspondence is upper hemicontinuous. Since the set of maximizers belongs to a finite set (the integers between 1 and $n - K(\mathbf{n})$), this implies that there exists $\delta^{\mathbf{n}}$ such that

$$\mu(\mathbf{n}, \delta^q) \subseteq \mu(\mathbf{n}, 1) \qquad \text{for all } \delta^q \geq \delta^{\mathbf{n}}.$$

Since the number of players is finite, $\mathscr{F}$ is a finite set. Therefore, we can find δ^* such that, for *every* $\mathbf{n} \in \mathscr{F}$,

$$\mu(\mathbf{n}, \delta^q) \subseteq \mu(\mathbf{n}, 1) \qquad \text{for all } \delta^q \geq \delta^*.$$

Now observe that $\mu(\mathbf{n}, 1)$ is precisely the set of integers that maximize (1) (recall Step 3 of the algorithm). This means that if $\delta \geq \delta^*$, then for every $\mathbf{n}$, if t^* maximizes the expression in (19) (i.e., $t^* \in \mu(\mathbf{n}, \delta)$), then

$$\frac{v(t^*, c(\mathbf{n} \cdot t^*))}{t^*} = \frac{v(t(\mathbf{n}), c(\mathbf{n} \cdot t(\mathbf{n})))}{t(\mathbf{n})} \equiv a(\mathbf{n}). \tag{20}$$

That is, t^* maximizes the expression in (1) as well.

It remains to be shown that if $\delta \geq \delta^*$, then $\mu(\mathbf{n}, \delta)$ contains only one such t^*, and that it is the *largest* integer maximizing "average worth" in (1), i.e., $\mu(\mathbf{n}, \delta) = \{t(\mathbf{n})\}$. Suppose not. From the construction of $t(\mathbf{n})$, this means that for some $\delta \geq \delta^*$ and for some $t^* \in \mu(\mathbf{n}, \delta)$, we have $t^* < t(\mathbf{n}) \equiv \hat{t}$. Therefore

$$\frac{1 - \delta}{\hat{t}} + \delta < \frac{1 - \delta}{t^*} + \delta.$$

Using this and the fact that $a(\mathbf{n}) > 0$, we see that

$$\frac{\hat{t}a(\mathbf{n})}{1 + \delta[\hat{t} - 1]} > \frac{t^* a(\mathbf{n})}{1 + \delta[t^* - 1]},$$

and recalling the definition of $a(\mathbf{n})$ from (20), we may conclude that

$$\frac{v(\hat{t}, c(\mathbf{n} \cdot \hat{t}))}{1 + \delta[\hat{t} - 1]} > \frac{v(t^*, c(\mathbf{n} \cdot t^*))}{1 + \delta[t^* - 1]}.$$

But this contradicts the fact that t^* maximizes the expression in (19). ■

We now establish a lemma that is useful for proving subsequent results.

LEMMA 5.2. *Consider the stage in which π has left the game and S is the set of active players. Let $\mathbf{n} \in \mathscr{F}$ denote the numerical coalition structure corresponding to π, and let $(x_i)_{i \in S}$ denote the equilibrium payoffs to each* *active player if he is the proposer at this stage. Suppose that for any* $t \in \{1, \dots, n - K(\mathbf{n})\}$ *the numerical coalition structure following* $(\mathbf{n} \cdot t)$ *is* $c(\mathbf{n} \cdot t)$. *Then, if* i *makes an acceptable proposal to coalition* T^* *with positive probability,*

(1) $j \in T^*$, $j \neq i$, *and* $x_k < x_j$ *implies* $k \in T^*$,

(2) $x_i \leq x_k$ *for all* $k \in S$.

Proof. Since i makes an acceptable proposal to T^* with positive probability and the resulting coalition structure is $c(\mathbf{n} \cdot t^*)$, it follows that

$$\begin{aligned} x_i &= v(t^*, c(\mathbf{n} \cdot t^*)) + \delta \sum_{j \in T^*;\, j \neq i} x_j \\ &\geq \max_{T \subseteq S;\, i \in T} \left[v(t, c(\mathbf{n} \cdot t)) - \delta \sum_{j \in T;\, j \neq i} x_j \right]. \end{aligned} \tag{21}$$

Part (1) is an immediate consequence of (21). Suppose part (2) is false, i.e., $x_k < x_i$ for some $k \in S$. Using the hypothesis that the coalition structure following $(\mathbf{n} \cdot t)$ is $c(\mathbf{n} \cdot t)$, it follows that if $k \notin T^*$, then k can form the coalition $(T^* \setminus \{i\}) \cup \{k\}$ and receive the same as x_i, a contradiction. Suppose $k \in T^*$. Then

$$\begin{aligned} x_k &\geq v(t^*, c(\mathbf{n} \cdot t^*)) - \delta \sum_{j \in T^*;\, j \neq k} x_j \\ &= v(t^*, c(\mathbf{n} \cdot t^*)) - \delta \sum_{j \in T^*;\, j \neq i} x_j + \delta x_k - \delta x_i, \end{aligned}$$

which implies, using (21), that $x_k \geq x_i$, but this is a contradiction. ■

PROOF OF THEOREM 3.1. Fix an equilibrium as described in the statement of the theorem, and let $\delta \in (\delta^*, 1)$, with δ^* as in Lemma 5.1. We proceed by induction on the cardinality of the set of active players, following the departure of any collection of players. If there is one active player left, then there is nothing to prove. Inductively, suppose that the theorem is valid at every stage with $K(\mathbf{n}(\pi)) = M + 1, \dots, n - 1$ for some $m \geq 0$.

Consider, now, a stage with $K(\mathbf{n}(\pi)) = m$. Let S be the set of active players, and let $\{x_j\}_{j \in S}$ denote the vector of equilibrium payoffs to player j if j is the proposer at this stage. Let T^* be a coalition that forms at this stage (with cardinality t^*), and let k be the proposer. We need to prove that

$$t^* = t(\mathbf{n}(\pi)). \tag{22}$$

Since every player in S makes an acceptable proposal to some coalition with positive probability, it follows immediately from the induction hypothesis and Part (2) of Lemma 5.2 that $x_j = x_i = x$ for all $i, j \in S$. It follows from the induction and the optimality of the proposal that

$$x = v(t^*, c(\mathbf{n}(\pi) \cdot t^*)) - \delta(t^* - 1)x \geq v(t, c(\mathbf{n}(\pi) \cdot t)) - \delta(t - 1)x$$

for all $t \in \{1, \dots, n - K(\mathbf{n}(\pi))\}$.

But this observation implies that x must also be the maximum value of the expression in (19), as t varies over the set $\{1, \dots, n - K(\mathbf{n}(\pi))\}$. Using Lemma 5.1, we may conclude that $t^* = \mathrm{t}(\mathbf{n})$. Of course, the payoff to a proposer is $a(\mathbf{n}, \delta)$. ■

LEMMA 5.3. *There exists $\hat{\delta} \in (\delta^*, 1)$ such that for all $\mathbf{n} \in \mathcal{F}$ and positive integers $t_1, \dots, t_k$ with $\mathbf{n} \cdot t_1 \cdots t_k \in \mathcal{F}$, the relationship*

$$a(\mathbf{n}) \geq a(\mathbf{n} \cdot t_1 \cdots t_k) \tag{23}$$

implies the relationship

$$a(\mathbf{n}, \delta) > \delta a(\mathbf{n} \cdot t_1 \cdots t_k, \delta) \qquad \text{for all } \delta \in (\hat{\delta}, 1), \tag{24}$$

where $a(\mathbf{n}, \delta)$, it will be recalled, is defined in (3).

Proof. Observe from (20) and (3) that, for each $\mathbf{n} \in \mathcal{F}$, $a(\mathbf{n}, \delta) \to a(\mathbf{n})$ as $\delta \to 1$. It follows that for each $\mathbf{n}$ such that *strict* inequality holds in (23), there exists $\delta^{\mathbf{n}} \in (0, 1)$ such that for all $\delta \in (\delta^{\mathbf{n}}, 1)$, (24) holds with a strict inequality. We focus on the case where equality holds in (23).

We proceed by differentiating both sides of (24) and examining their comparative magnitudes at $\delta = 1$. If we can show that

$$\left.\frac{d}{d\delta} a(\mathbf{n}, \delta)\right|_{\delta = 1} < \left.\frac{d}{d\delta} \delta a(\mathbf{n} \cdot t_1 \cdots t_k, \delta)\right|_{\delta = 1}, \tag{25}$$

then, in light of the fact that equality holds in (23), we will be able to conclude that there exists $\delta^{\mathbf{n}} \in (0, 1)$ such that for all $\delta \in (\delta^{\mathbf{n}}, 1)$, (24) holds with strict inequality. The proof of the lemma is then complete by noting that $\mathcal{F}$ is a finite set, so that the required $\hat{\delta}$ can be obtained by choosing the maximum of the values $\delta^{\mathbf{n}}$ over $\mathbf{n}$, and δ^*.

It remains, then, to establish (25) in the case where (23) holds with equality. To simplify the notation, let $t \equiv t(\mathbf{n})$, $t' \equiv t(\mathbf{n} \cdot t_1 \cdots t_k)$, and $a \equiv a(\mathbf{n}) = a(\mathbf{n} \cdot t_1 \cdots t_k) > 0$ (note that $K(\mathbf{n}) < n - 1$ since $\mathbf{n} \cdot t_1 \cdots t_k \in \mathcal{F}$). Then, recalling the definition of $a(\mathbf{n})$ from (20) and noting that $c(\mathbf{n}) = c(\mathbf{n} \cdot t(\mathbf{n}))$, it is clear that

$$a(\mathbf{n}, \delta) = \frac{at}{1 + \delta(t - 1)},$$

while

$$\delta a(\mathbf{n} \cdot t_1 \ \cdots \ t_k, \delta) = \frac{\delta a t'}{1 + \delta(t' - 1)}.$$

Simple computation now reveals that

$$\left. \frac{d}{d\delta} a(\mathbf{n}, \delta) \right|_{\delta = 1} = -\frac{a(t-1)}{t},$$

while

$$\left. \frac{d}{d\delta} \delta a(\mathbf{n} \cdot t_1 \ \cdots \ t_k, \delta) \right|_{\delta = 1} = a - \frac{a(t'-1)}{t'}.$$

Because $a > 0$, it follows right away from these two expressions that (25) must hold.

The proof of the lemma is completed, as already described, by letting $\hat{\delta} = \max\{\delta^*, \max_{\mathbf{n} \in \mathscr{F}} \delta^n\}$. ■

Proof of Theorem 3.2. Assume (4). Pick any $\delta \in (\hat{\delta}, 1)$, where $\hat{\delta}$ is given by Lemma 5.3.

Consider any stationary strategy σ as follows: in every subgame following the departure of π, player i makes a proposal to a coalition of size $t(\mathbf{n}(\pi))$. He offers to each partner a payoff $\delta a(\mathbf{n}, \delta)$ in the event that the numerical coalition structure $c(\mathbf{n})$ is formed, and any other payoff division otherwise. All such offers are accepted by respondents (other responses are described in the obvious way; for a description, see (ii) and (iii) in the proof of Theorem 2.1). We will show that σ is an equilibrium.

To this end, consider any stage described by π. Along the proposed strategy profile σ, a proposer receives $a(\mathbf{n}(\pi), \delta)$. Therefore, the only way that a proposer can possibly deviate gainfully is by making an unacceptable proposal. Given the strategies of the other players, this will result in the formation of coalitions of cardinalities $t(\mathbf{n})$, $t(\mathbf{n} \cdot t(\mathbf{n}))$, and so on. Thus the deviant proposer will ultimately receive a payoff that is is bounded above by $\delta a(\mathbf{n} \cdot t_1 \ \cdots \ t_k, \delta)$, where $t_1 \ \cdots \ t_k$ is a finite string of the form $t(\mathbf{n}) \cdot t(\mathbf{n} \cdot t(\mathbf{n})) \cdots$. Applying (4) repeatedly, we see that

$$a(\mathbf{n}) \geq a(\mathbf{n} \cdot t_1 \ \cdots \ t_k).$$

But, then, by Lemma 5.3 and the fact that $\delta > \hat{\delta}$, we conclude that (24) holds. This means that the derivation cannot be profitable.

It is now easy to see that as a responder, a player cannot gainfully deviate from σ. Consequently, σ is an equilibrium. ■

Proof of Theorem 3.3. Suppose, on the contrary, that there is $\hat{\delta} \in (0, 1)$ such that for all $\delta \in (\hat{\delta}, 1)$ there exists a pure-strategy no-delay equilibrium but (4) fails. This means that there exists $\mathbf{n} \in \mathscr{F}$ such that $\mathbf{n} \cdot t(\mathbf{n}) \in \mathscr{F}$ as well, and such that $a(\mathbf{n} \cdot t(\mathbf{n})) > a(\mathbf{n})$. It follows that there exists $\delta \in (0, 1)$ such that

$$\delta a(\mathbf{n} \cdot t(\mathbf{n}), \delta) > a(\mathbf{n}, \delta) \qquad \text{for all } \delta \in (\bar{\delta}, 1). \tag{26}$$

Consider any $\delta > \max\{\hat{\delta}, \bar{\delta}\}$, and fix a pure-strategy no-delay equilibrium σ. Consider any subgame where π has left, where $\mathbf{n}(\pi) = \mathbf{n}$. Let i be the first proposer in this subgame. Since σ is a pure-strategy no-delay equilibrium and $\delta \geq \hat{\delta} \geq \delta^*$, i makes an acceptable proposal to some determinate coalition of size $t(\mathbf{n})$. Because $\mathbf{n} \cdot t(\mathbf{n}) \in \mathscr{F}$, there must exist a player j who is not included in the proposal by player i, and thereafter picks up a present value of $a(\mathbf{n} \cdot t(\mathbf{n}), \delta)$ in the very next stage.

Now consider another subgame (in the same stage) so that exactly the same set of players has left (and in the same structure), but j is the first proposer instead of i. Because σ is a pure-strategy no-delay equilibrium, j is also supposed to make an acceptable proposal to a coalition of size $t(\mathbf{n})$, picking up a $(\mathbf{n}, \delta)$. However, suppose that he deviates by making an unacceptable offer to i. By stationarity, we are then in the precise situation of the preceding paragraph, with a delay of one unit of time. Thus, by making an unacceptable proposal to player i, j receives a present value of $\delta a(\mathbf{n} \cdot t(\mathbf{n}), \delta)$. By (26) this deviation is profitable. This contradicts the fact that we have an equilibrium. ■

Proof of Theorem 3.4. Fix $\hat{\delta}$ as given in Lemma 5.3, and any equilibrium. We will show that it must be no-delay. The proof is by induction on the cardinality of the set of active players. At every stage when there is only one active player left, the subgame equilibrium is trivially no-delay. Now suppose that for any π such that $K(\mathbf{n}(\pi)) \geq m + 1, \dots, n - 1$, for some $m \geq 0$, the subgame equilibrium is no-delay. Consider a stage described by a structure of departed players, π, with the property that $K(\mathbf{n}(\pi)) = m$. Let $\mathbf{n} \equiv \mathbf{n}(\pi)$. Let S be the set of active players. Let $\{x_i\}_{i \in S}$ denote the vector of equilibrium payoffs to each player, if he is the proposer at this stage. Without loss of generality, number of players such that $x_1 \leq \cdots \leq x_s$.

Because of the regularity condition that guarantees $a(\mathbf{n}) > 0$, *some* player must make an acceptable proposal with positive probability. From Lemma 5.2 it follows that there is no loss of generality in assuming that player 1 does so to coalition $T^* = \{1, \dots, t^*\}$.

Since player 1 makes an acceptable proposal to T^*, it must be the case that, for any $t \leq n - m$,

$$x_1 = v(t^*, c(\mathbf{n} \cdot t^*)) - \delta \sum_{j=1}^{t^*} x_j$$

$$\geq v(t, c(\mathbf{n} \cdot t)) - \delta \sum_{j=2}^{t} x_j \qquad \text{for all } j \in S. \tag{27}$$

We claim that

$$\frac{v(t, c(\mathbf{n} \cdot t))}{t} \leq \frac{v(t^*, c(\mathbf{n} \cdot t^*))}{t^*} \qquad \text{for all } t \leq t^*. \tag{28}$$

Suppose not. Then there exists $\hat{t} < t^*$ such that

$$\frac{v(\hat{t}, c(\mathbf{n} \cdot t))}{\hat{t}} > \frac{v(t^*, c(\mathbf{n} \cdot t^*))}{t^*}.$$

From Lemma 5.3, and given our choice of δ, it follows that

$$\hat{a} \equiv \frac{v(\hat{t}, c(\mathbf{n} \cdot \hat{t}))}{1 + \delta(\hat{t} - 1)} > a^* \equiv \frac{v(t^*, c(\mathbf{n}, t^*))}{1 + \delta(t^* - 1)}$$

or

$$v(\hat{t}, c(\mathbf{n} \cdot \hat{t})) - \delta(\hat{t} - 1)\hat{a} > v(t^*, c(\mathbf{n} \cdot t^*)) - \delta(t^* - 1)a^*.$$

Since $\hat{a} > a^*$, we may combine this inequality with (27) and rearrange terms to see that

$$(t^* - \hat{t})\hat{a} > \sum_{j=\hat{t}+1}^{t^*} x_j,$$

which permits us to conclude that

$$\hat{a} > x_j \qquad \text{for all } j = 1, \ldots, \hat{t}. \tag{29}$$

However, (29) implies that

$$v(\hat{t}, c(\mathbf{n} \cdot \hat{t})) - \delta \sum_{j=2}^{\hat{t}} x_j > v(\hat{t}, c(\mathbf{n} \cdot \hat{t})) - \delta(\hat{t} - 1)\hat{a} = \hat{a} > x_1,$$

a contradiction to (27). This completes the proof of (28).

Suppose, now, that the theorem is false; i.e., there exists a player who makes an unacceptable offer. By Lemma 5.2, there is no loss of generality, in assuming that this is player s. Notice that $s \in T^*$, otherwise his expected payoff would be δx_s rather than x_s. By the induction hypothesis, the coalition structure following t^* is $c(\mathbf{n} \cdot t^*) = (n \cdot t^* \cdot t_2 \cdots t_m)$. Suppose s belongs to T_k, where $2 \leq k \leq m$. Applying the induction hypothesis again,

$$x_s \leq \delta a(\mathbf{n} \cdot t^* \cdots t_{k-1}, \delta). \tag{30}$$

Using (28) and condition (6), it follows that $a(\mathbf{n} \cdot t^* \cdots t_{k-1}) \leq a(\mathbf{n})$. Combining this observation with Lemma 5.3 and (30), we conclude that

$$x_s \leq \delta a(\mathbf{n} \cdot t \cdots t_{k-1}, \delta) < a(\mathbf{n}, \delta) = a. \tag{31}$$

Since $a + \delta(t(\mathbf{n}) - 1)a = v(t(\mathbf{n}), c(\mathbf{n}))$, (31) implies that

$$v(t(\mathbf{n}), c(\mathbf{n})) - \delta \sum_{j=2}^{t(\mathbf{n})} x_j > x_1,$$

which contradicts (27). ■

Proof of Theorem 3.5. Suppose the theorem is false. Proceed, by induction, exactly as in the proof of Theorem 3.4, using the same notation, leading up to condition (30), i.e.,

$$x_1 \leq \cdots \leq x_s \leq \delta a(\mathbf{n} \cdot t^* \cdots t_{k-1}, \delta).$$

Of course $\delta t_k a(\mathbf{n} \cdot t^* \cdots t_{k-1}, \delta) < v(t_k, c(\mathbf{n} \cdot t^*))$, which implies that

$$\delta t_k x_s < v(t_k, c(\mathbf{n} \cdot t^*)). \tag{32}$$

By superadditivity, we know that a coalition of $t^* + t_k$ can obtain at least the sum of the worths of coalitions t^* and t_k, i.e.,

$$v(t^* + t_k, c(\mathbf{n} \cdot t^* + t_k)) \geq v(t^*, c(\mathbf{n} \cdot t^*)) + v(t_k, c(\mathbf{n} \cdot t^*)),$$

which means that

$$\begin{aligned} & v(t^* + t_k, c(\mathbf{n} \cdot t^* + t_k)) - \delta \sum_{j=2}^{t^*+t_k} x_j \\ & \quad \geq \left[v(t^*, c(\mathbf{n} \cdot t^*)) - \delta \sum_{j=2}^{t^*} x_j \right] + \left[v(t_k, c(\mathbf{n} \cdot t^*)) - \delta \sum_{j=t^*+1}^{t^*+t_k} x_j \right]. \end{aligned}$$

Since $x_j \le x_s$ for all j, it follows from (32) that the last term is positive. But, then, we have

$$v(t^* + t_k, c(\mathbf{n} \cdot t^* + t_k)) - \delta \sum_{j=2}^{t^*+t_k} x_j > v(t^*, c(\mathbf{n} \cdot t^*)) - \delta \sum_{j=2}^{t^*} x_j,$$

which implies that player 1 receives more than x_1 by making an acceptable offer to players $2, \dots, t^* + t_k$, a contradiction. ∎

Proof of Theorem 3.6. We begin with a description of the function $t(\cdot)$. For each $\mathbf{n} \in \mathcal{F}$, let $R(\mathbf{n})$ denote the number of coalitions in $\mathbf{n}$ and let $m(\mathbf{n}) \equiv n - K(\mathbf{n})$. Following the arguments in Bloch (1996) it can be shown that

$$t(\mathbf{n}) = \begin{cases} 1 & \text{if } m(\mathbf{n}) < (R(\mathbf{n}) + 1)^2, \\ m(\mathbf{n}) & \text{if } (R(\mathbf{n}) + 1)^2 \le m(\mathbf{n}) < (R(\mathbf{n}) + 2)^2 + 1, \\ 1 & \text{if } m(\mathbf{n}) \ge (R(\mathbf{n}) + 2)^2 + 1. \end{cases} \tag{33}$$

Given this description of $t(\cdot)$, we first verify that $\mathbf{n}^*$ is of the form described in the statement of the theorem. Starting at ϕ, $t(\cdot)$ dictates that singletons must form (so that $t(\mathbf{n}) = 1$ and $R(\mathbf{n})$ equals the number of elements of $\mathbf{n}$ for all such $\mathbf{n}$) until we reach the first nonnegative integer L such that

$$n - L < (L + 2)^2 + 1. \tag{34}$$

Given (33), it remains to show that L also satisfies the inequality

$$n - L \ge (L + 1)^2.$$

Suppose not. Then $n - L < (L + 1)^2$. This means that L is a positive integer, so that $L' \equiv L - 1$ is a nonnegative integer. But, then, $n - L' > (L' + 1)^2 + 2$, which contradicts the definition of L in (34).

To complete the proof, we verify that condition (6) holds. To do so, we note that starting from any $\mathbf{n}$, (33) guarantees that larger coalitions always form later than smaller coalitions. It immediately follows that (4) of Theorem 3.2 is met.

Now suppose $t(\mathbf{n}) = 1$. Then it is clear that $t_l(\mathbf{n}) = 1$ and (6) is the same as (4) of Theorem 3.2, which we know is satisfied.

It remains to consider the case in which $t(\mathbf{n}) = m(\mathbf{n}) > 1$. In this case, we may conclude from (33) that $m(\mathbf{n}) < (R(\mathbf{n}) + 2)^2 + 1$. Because (as a result of this inequality) we have $m(\mathbf{n}) - 1 < (\{R(\mathbf{n}) + 1\} + 1)^2$, we may infer that, for any $\mathbf{n}'$ with $\mathbf{n} \subset \mathbf{n}'$, $t(\mathbf{n}') = 1$.

It follows that, for any alternative choice $t' < t(\mathbf{n})$,

$$c(\mathbf{n} \cdot t') = (\mathbf{n} \cdot t' \cdot 1 \cdot \cdots \cdot 1).$$

Moreover, the same logic tells us that if we consider the structure $\mathbf{n}''$ derived from $c(\mathbf{n} \cdot t')$ with t' removed, $t(\mathbf{n}'') = 1$ as well. This implies, in particular, that, for all $t' < t(\mathbf{n})$,

$$\frac{v(t', c(\mathbf{n} \cdot t'))}{t'} = \frac{v(t', (\mathbf{n} \cdot t' \cdot 1 \cdot \cdots \cdot 1))}{t'}$$
$$< v(1, (\mathbf{n} \cdot 1 \cdot \cdots \cdot 1)) = v(1, c(\mathbf{n} \cdot 1)) \tag{35}$$

and this tells us right away that $t_l(\mathbf{n}) = 1$ for any $l < t(\mathbf{n})$. Equation (35) also contains the information that

$$a(\mathbf{n} \cdot t_l(\mathbf{n})) = v(1, c(\mathbf{n} \cdot 1)). \tag{36}$$

On the other hand, we know that

$$a(\mathbf{n}) = \frac{v(t, c(\mathbf{n} \cdot t))}{t} \geq v(1, c(\mathbf{n} \cdot 1)). \tag{37}$$

Combining (36) and (37), we see that (6) is verified. The result that $\mathbf{n}^*$ is the unique numerical coalition structure now follows from Theorem 3.4. ■

PROOF OF THEOREM 4.1. The proof is broken up into several steps. Fix any characteristic function w on a player set S and recall the construction in Section 4.1.1 (see (7)–(12)). Our first observation is

LEMMA 5.4. *$A^k \geq A^{k+1}$ for all k.*

Proof. Suppose not. Consider the first integer k for which the inequality is violated. Using (10), we see that if $T = T_1 \cup \cdots \cup T_{k+1}$ solves the maximization problem there, then

$$t_{k+1}A^{k+1} + t_k A^k = w(T) - \sum_{j=1}^{k-1} A^j t_j$$

(where the empty sum by convention is assumed to give zero value). But because $A^k < A^{k+1}$ and $t_{k+1} > 0$, this implies that

$$A^k < \frac{w(T) - \sum_{j=1}^{k-1} A^j t_j}{t_k + t_{k+1}},$$

which contradicts the construction of A^k, since T is certainly an admissible set of the maximization problem defining A^k. ■

LEMMA 5.5. *For any characteristic function w on a player set S, $\{a_i(w)\}_{i \in S}$ is a solution to the following two requirements:*

$$a_i(w) = \max_{T \subseteq S;\, i \in T} \left[v(T) - \sum_{j \in T \setminus \{i\}} a_j(w) \right]. \tag{38}$$

and, for every $i \in S$,

$$\text{For some } T_i \text{ that solves } (38),\ a_j(w) \geq a_i(w) \qquad \text{for all } j \in T_i. \tag{39}$$

Moreover, there are no other solutions satisfying (38) *and* (39).

Proof. Use Lemma 5.4 along with the definitions of $\{a_i(w)\}$ (given by (8) and (11)) to see that

$$a_i(w) = \max \left[v(T) - \sum_{j \in T \setminus \{i\}} a_j(w) \right], \tag{40}$$

where the maximum is taken over *only those sets* T such that $T \subseteq S$, $i \in T$, and $j \in T$ implies $a_j(w) \geq a_i(w)$ for all i. We need the stronger implication (38), which imposes less restrictions on the maximizing set T. To see that this is automatically implied, suppose, on the contrary, that there is a coalition T such that

$$a_i(w) < v(T) - \sum_{j \in T \setminus \{i\}} a_j(w). \tag{41}$$

Given the property (40), this can only be the case if, for some $y \in T$, $a_j(w) < a_i(w)$. Let k be the index in T such that $a_k(w)$ is the *smallest*. Then, rearranging (41), we see that

$$a_k(w) < v(T) + \sum_{j \in T \setminus \{k\}} a_j(w).$$

But this contradicts the property (40) for the index k, because $a_j(w) \geq a_k(w)$ for all $j \in T$. So (28) is established, and (40) assumes us that (39) is satisfied as well.

Finally, we must show that there are no other solutions to (38) and (39). We adapt an argument from Chatterjee *et al.* (1993, proof of Prop. 1). Suppose, contrary to the claim, that there are distinct solutions $\{a_i\}$ and $\{b_i\}$ to (38) and (39). Let $\Omega \equiv \{i \in S \mid a_i \neq b_i\}$. Choose $k \in \Omega$ such that

(without loss of generality)

$$a_k = \max\{z \mid z = a_i \text{ or } z = b_i \text{ for } i \in \Omega\}. \tag{42}$$

Using (39), we may pick a coalition T_k such that

$$a_k = v(T_k) - \sum_{j \in T_k \setminus \{k\}} a_j \tag{43}$$

and such that

$$a_j \geq a_k \qquad \text{for every } j \in T_k. \tag{44}$$

Because $\{b_i\}$ satisfies (38), we have

$$b_k \geq v(T_k) - \sum_{j \in T_k \setminus \{k\}} b_j. \tag{45}$$

Because of (42) and (44), it must be the case that $b_j \leq a_j$ for all $j \in T_k$. But, then, combining this information with (43) and (45), we see that $b_k \geq a_k$. This contradicts (42). ■

LEMMA 5.6. *For each characteristic function w with player set S and each $\delta \in (0, 1)$, there exists a unique vector of numbers $\{a_i(w, \delta)\}_{i \in S}$ such that, for every $i \in S$,*

$$a_i(w, \delta) = \sum_{T \subseteq S;\, i \in T} \left[v(T) - \delta \sum_{j \in T \setminus \{i\}} a_j(w, \delta) \right]. \tag{46}$$

Moreover,

$$\lim_{\delta \to 1} a_i(w, \delta) = a_i(w), \tag{47}$$

where $\{a_i(w)\}_{i \in S}$ is defined in (8) *and* (11).

Proof. The first part of the lemma, which establishes the uniqueness of the vector $\{a_i(w, \delta)\}_{i \in S}$, is proved in Chatterjee *et al.* (1993, Prop. 1). To prove the limit result asserted in the second part, consider any limit point $\{a_i\}$ of $\{a_i(w, \delta)\}$. Choose a subsequence of δ such that $a_i(w, \delta) \to a_i$ for all i as $\delta \to 1$ along this subsequence. By passing to the limit in (46), we see that $\{a_i\}$ must satisfy (38).

Next, consider any coalition $T_i(\delta)$ that attains the maximum in (46). By Lemma 2 in Chatterjee *et al.* (1993), we have $a_j(w, \delta) \geq a_i(w, \delta)$ for all $j \in T_i(\delta)$. Let T_i be some limit point of $T_i(\delta)$ as $\delta \to 1$ along the subsequence of the previous paragraph. Then T_i solves the maximization problem implicit in (38). And, certainly, $a_j \geq a_i$ for all $j \in T_i$. We have therefore proved that *every* limit point of $\{a_i(w, \delta)\}$ (as $\delta \to 1$) must satisfy (38) and (39).

However, Lemma 5.5 tells us that $\{a_i(w)\}$, as defined by (8) and (11), is the only solution to (38) and (39). ■

Our next step contains the heart of the argument. We first construct a finite collection of characteristic functions as follows. For each coalition T in the given player set, consider the finite set of real numbers $f(T)$ given by

$$f(T) \equiv \{v(T, \pi) \mid \pi \text{ is a coalition structure of } N \text{ and } T \in \pi\}.$$

Now let $\mathscr{W}$ be the collection of characteristic functions with the property that the player set is $S \subseteq N$ and, for each coalition T of S, $w(T) \in f(T)$. Observe that $\mathscr{W}$ is a finite set.

LEMMA 5.7. *Pick any characteristic function w in $\mathscr{W}$ (with player set S) and $\delta \in (0,1)$. By Lemma 5.6, there exists a unique vector $\{a_i(w, \delta)\}$ satisfying (46). For each i, define $\mathscr{C}_i(w, \delta)$ by the collection of coalitions that achieve the maximum in (46).*

Then there exists $\delta^* \in (0,1)$ such that, for all $\delta \in (\delta^*, 1)$ and all $w \in \mathscr{W}$,

$$\mathscr{C}_i(w, \delta) \subseteq \mathscr{C}_i(w), \tag{48}$$

where $\mathscr{C}_i(w)$ is defined in (9) and (12).

Proof. Pick any characteristic function w in $\mathscr{W}$ (with player set S). Define $\mathscr{C}_i(w, \delta)$ for each i as in the statement of the lemma. Because the set $\mathscr{W}$ is finite, it will be sufficient to show that there exists a threshold $\delta(w) \in (0,1)$ such that, for all $\delta \in (\delta(w), 1)$, $\mathscr{C}_i(w, \delta) \subseteq \mathscr{C}_i(w)$. The result then follows by considering the maximum threshold $\delta(w)$, over $w \in \mathscr{W}$.

Suppose, on the contrary, that there is some i, a subsequence δ^m converging to 1, and a coalition T such that $T \in \mathscr{C}_i(w, \delta^m)$ for all m, but $T \notin \mathscr{C}_i(w)$. Let $i \in U_{k+1}$ for some $k \geq 0$ (see the construction in Section 4.1.1). Indeed, take the smallest value of k for which this is so.

Now, there are two cases to consider.

Case 1. T is not a solution to the maximization problem defining A^{k+1} in (10). Write T in the form $T_1 \cap \cdots \cap T_{k+1}$, where $T_l \subseteq U_l$ for all $l = 1, \ldots, k$, and $T_{k+1} \subseteq S \setminus \bigcup_{l=1}^{k} U_l$. Note that T_{k+1} is nonempty, because $i \in T$ and $i \in U_{k+1}$ by assumption. It follows that

$$A^{k+1} \geq \frac{w(T) - \sum_{l=1}^{k} A^l t_l}{t_{k+1}},$$

or, equivalently, that

$$A^{k+1} > w(T) - \sum_{l=1}^{k} A^l t_l - (t_{k+1} - 1) A^{k+1}. \tag{49}$$

Because $T \in \mathscr{C}_i(w, \delta^m)$ for all m, we have

$$a_i(w, \delta^m) = w(T) - \delta^m \sum_{j \in T \setminus \{i\}} a_j(w, \delta^m)$$

for all m. Using the decomposition $T_1 \cup \cdots \cup T_{k+1}$ introduced above, this is the same as stating that

$$a_i(w, \delta^m) = w(T) - \delta^m \left[\sum_{l=1}^{k} \sum_{j \in T_l} a_j(w, \delta^m) + \sum_{j \in T_{k+1} \setminus \{i\}} a_j(w, \delta^m) \right]. \tag{50}$$

Now we make four observations. First, note that by Lemma 5.6 (see (47)), $\lim_{m \to \infty} a_j(w, \delta^m) = a_j(w)$ for all $j \in S$. Second, if $1 \leq l \leq k$ and $j \in T_l$, then $a_j(w) = A^l$ by construction. Third, because $i \in U_{k+1}$, it is also true that $a_i(w) = A^{k+1}$. Finally, by Lemma 2 in Chatterjee *et al.* (1993), $a_j(w, \delta^m) \geq a_i(w, \delta^m)$ for all $j \in T_{k+1} \setminus \{i\}$. So $\lim_{m \to \infty} a_j(w, \delta^m) \geq \lim_{m \to \infty} a_i(w, \delta) = A^{k+1}$. However, this last inequality cannot hold strictly, because $T_{k+1} \subseteq S \setminus \bigcup_{l=1}^{k} U_l$, and because of Lemma 5.4. Thus $\lim_{m \to \infty} a_j(w, \delta^m) = A^{k+1}$ for all $j \in T_{k+1} \setminus \{i\}$. Combining these four observations and passing to the limit as $m \to \infty$ in (50), we conclude that

$$A^{k+1} = w(T) - \sum_{l=1}^{k} A^l t_l - (t_{k+1} - 1) A^{k+1}.$$

But this last equality directly contradicts (49). This means that Case 1 is impossible. This leaves as the only remaining possibility

Case 2. $T = T_1 \cup \cdots \cup T_{k+1}$ is a solution to the maximization problem defining A^{k+1} in (10), but fails to minimize $\Delta(T_1 \cup \cdots \cup T_{k+1})$.

For any function f: $(0, 1] \mapsto R$, define $h(f(\delta)) \equiv (f(\delta) - f(1))/(\delta - 1)$ for $\delta \in (0, 1)$. If f possesses a left-hand derivative at $\delta = 1$, denote this by $f' \equiv \lim_{\delta \to 1} h(f(\delta))$. We make the following

CLAIM. *Suppose that there exists some* $\bar{\delta} \in (0, 1)$ *such that, for all* $\delta \in (\bar{\delta}, 1)$ *and for all* $i \in \bigcup_{s=1}^{k} U_s$, $\mathscr{C}_i(w, \delta) \subseteq \mathscr{C}_i(w)$. *Then, for all such* i, $a_i(w, \delta)$ *has a left-hand derivative at* $\delta = 1$, $a_i'(w) = \Delta^s$, *where* $i \in U_s$.

Proof. The proof is by induction on s. Begin with the inductive step, assuming that the lemma is true for all indices $s = 1, \ldots, l$, for some $l < k$. Pick any $i \in U_{l+1}$. Suppose $\delta \geq \bar{\delta}$. Pick some sequence $\delta \uparrow 1$ and a coalition $T \in \mathscr{C}_i(w, \delta) \subseteq \mathscr{C}_i(w)$ along this sequence. Let T be the form $T_1 \cup \cdots \cup T_{l+1}$, where $T_s \subseteq U_s$ for all $s = 1, \ldots, l+1$, and $i \in T_{l+1}$. Then

$$a_i(w, \delta) = w(T) - \delta \sum_{s=1}^{l} \sum_{j \in T_s} a_j(w, \delta) - \delta \sum_{j \in T_{l+1} \setminus \{i\}} a_j(w, \delta).$$

This can be rewritten as

$$\sum_{j \in T_{l+1}} a_j(w, \delta) = w(T) - \delta \sum_{s=1}^{l} \sum_{j \in T_s} a_j(w, \delta) - (\delta - 1) \sum_{j \in T_{l+1} \setminus \{i\}} a_j(w, \delta).$$

Subtracting from both sides of this equation the corresponding expression evaluated at $\delta = 1$ and using the fact that $a_i(w, 1) = A^s$ for $i \in T_s$, we get

$$\sum_{j \in T_{l+1}} \left[a_j(w, \delta) - A^{l+1} \right] = - \left[\sum_{s=1}^{l} \sum_{j \in T_s} a_j(w, \delta) - \sum_{s=1}^{l} t_s A^s \right] - (\delta - 1) \sum_{s=1}^{l} \sum_{j \in T_s} a_j(w, \delta) - (\delta - 1) \sum_{j \in T_{l+1} \setminus \{i\}} a_j(w, \delta).$$

Dividing both sides by $(\delta - 1) t_{l-1}$, we have

$$\frac{h\left(\sum_{j \in T_{l+1}} a_j(w, \delta)\right)}{t_{l+1}} = - \frac{h\left(\sum_{s=1}^{l} \sum_{j \in T_s} a_j(w, \delta)\right)}{t_{l+1}} - \frac{\sum_{s=1}^{l} \sum_{j \in T_s} a_j(w, \delta)}{t_{l+1}} - \frac{\sum_{j \in T_{l+1} \setminus \{i\}} a_j(w, \delta)}{t_{l+1}}. \tag{51}$$

By the induction hypothesis, the limit of the first term on the right-hand side of (51) is $-\sum_{s=1}^{l} t_s \Delta^s / t_{l+1}$. The limit of the second term is clearly $-\sum_{s=1}^{l} t_s A^s / t_{l+1}$ and that of the third term is $-(t_{l+1} - 1) A^{l+1} / t_{l+1}$.

Thus the limit, as $\delta \to 1$, of the left-hand side of (51) is well defined and

$$\lim_{\delta \to 1} \frac{h\left(\sum_{j \in T_{l+1}} a_j(w, \delta)\right)}{t_{l+1}} = -\frac{\sum_{s=1}^{l} t_s[A^s + \Delta^s] + (t_{l+1} - 1)A^{l+1}}{t_{l+1}} = \Delta^{l+1}. \tag{52}$$

By Lemma 1 of Chatterjee *et al.* (1993), $a_j(w, \delta) \geq a_i(w, \delta)$ for all δ and for all $j \in T_{l+1}$. Moreover, $a_i(w, \delta)$ and $a_j(w, \delta)$ converge to the same limit A^{l+1}. It follows that

$$h(a_i(w, \delta)) \geq \frac{h\left(\sum_{j \in T_{l+1}} a_j(w, \delta)\right)}{t_{l+1}} \qquad \text{for all } \delta < 1.$$

This, along with (52), yield

$$\liminf_{\delta \to 1} h(a_i(w, \delta)) \geq \lim_{\delta \to 1} \frac{h\left(\sum_{j \in T_{l+1}} a_j(w, \delta)\right)}{t_{l+1}} = \Delta^{l+1}. \tag{53}$$

Since $i \in U_{l+1}$ was arbitrary and $T_{l+1} \subseteq U_{l+1}$, we may conclude that (53) holds for every $j \in T_{l+1}$, i.e.,

$$\liminf_{\delta \to 1} h(a_j(w, \delta)) \geq \lim_{\delta \to 1} \frac{h\left(\sum_{j \in T_{l+1}} a_j(w, \delta)\right)}{t_{l+1}} = \Delta^{l+1} \qquad \text{for all } j \in T_{l+1}. \tag{54}$$

But this must mean that, for every $j \in T_{l+1}$, the left-hand side of (54) is the limit of $h(a_i(w, \delta))$ as $\delta \to 1$, i.e.,

$$a_i'(w) \equiv \lim_{\delta \to 1} h(a_i(w, \delta)) = \Delta^{l+1} \qquad \text{for all } j \in T_{l+1}.$$

This completes the inductive step of the proof.

The first step (which may be identified with $l = 0$ in the argument above) is proved in exactly the same way. Note that all sums of the form $\sum_{s=1}^{l}$ are 0, so that all reliance on induction can be dispensed with in this step. ∎

Now return to the main proof. Recall that we presumed that the desired result was false, and that k is the *smallest* index such that there is

$i \in U_{k+1}$ where the result fails. Therefore (in case $k \geq 1$), the conditions of the Claim apply, and we may take it that $a_i'(w) = \Delta^s$ for all $i \in U_s$ and $s \leq k$. If $k = 1$, then no such restriction is needed in the argument below.

Because T solves i's problem for δ^m, we see that

$$\begin{aligned} a_i(w, \delta^m) = w(T) &- \delta^m \sum_{s=1}^{k} \sum_{j \in T_s} a_j(w, \delta^m) \\ &- \delta^m \sum_{j \in T_{k+1} \setminus \{i\}} a_j(w, \delta^m). \end{aligned}$$

This can be written as

$$\begin{aligned} \sum_{j \in T_{k+1}} a_j(w, \delta^m) = w(T) &- \delta^m \sum_{s=1}^{k} \sum_{j \in T_s} a_j(w, \delta^m) \\ &-(\delta^m - 1) \sum_{j \in T_{k+1} \setminus \{i\}} a_j(w, \delta^m). \end{aligned}$$

By the Claim and the kind of argument used in its proof, it is now easy to see that

$$\begin{aligned} \lim_{\delta \to 1} & \frac{h\left(\sum_{j \in T_{k+1}} a_j(w, \delta)\right)}{t_{k+1}} \\ &= -\frac{\sum_{s=1}^{k} t_s[A^s + \Delta^s] + (t_{k+1} - 1)A^{k+1}}{t_{k+1}} \\ &= \Delta(T) \geq \Delta^{k+1}, \end{aligned}$$

where the last inequality holds by assumption, because we are in Case 2. Using the fact that $a_j(w, \delta^m) \geq a_i(w, \delta^m)$ for all $j \in T_{k+1} \setminus \{i\}$, we can assert that

$$\liminf_{\delta \to 1} h(a_i(w, \delta)) \geq \Delta(T) > \Delta^{k+1}.$$

As Case 1 has been shown to be impossible, a similar agreement can be used to show that

$$\liminf_{\delta \to 1} h\left(a_j(w, \delta)\right) > \Delta^{k+1} \qquad \text{for every } j \in U_{k+1}. \tag{55}$$

Now pick any coalition $T^* = T_1^* \cup \cdots \cup T_{k+1}^*$ in $\mathscr{C}_i(w)$. Define

$$b_i(w, \delta) \equiv w(T^*) - \delta \sum_{s=1}^{k} \sum_{j \in T_s^*} a_j(w, \delta) - \delta \sum_{j \in T_{k+1}^* \setminus \{i\}} a_j(w, \delta).$$

Clearly, for all m,

$$a_i(w, \delta^m) \geq b_i(w, \delta^m).$$

On the other hand, note that both $a_i(w, \delta^m)$ and $b_i(w, \delta)$ converge to the same limit A^{k+1}. Combining these two pieces of information, we may conclude that

$$\liminf_{\delta \to 1} h(b_i(w, \delta)) \geq \liminf_{\delta \to 1} h(a_i(w, \delta)). \tag{56}$$

Now return to the expression that defines $b_i(w, \delta)$, and construct the expressions used in defining the derivative, using the Claim and the fact that $a_i'(w) + \Delta^s$ for all $s \leq k$. This yields

$$\begin{aligned}
&\liminf_{\delta \to 1} h(b_i(w, \delta)) \\
&\qquad = -\sum_{s=1}^{k} t_s^*[A^s + \Delta^s] - (t_{k+1}^* - 1)A^{k+1} \\
&\qquad\quad - \limsup_{\delta \to 1} \sum_{j \in T_{k+1}^* \setminus \{i\}} h(a_j(w, \delta)) \\
&\qquad = t_{k+1}^* \Delta^{k+1} - \limsup_{\delta \to 1} \sum_{j \in T_{k+1}^* \setminus \{i\}} h(a_j(w, \delta)) \\
&\qquad < t_{k+1}^* \Delta^{k+1} - (t_{k+1}^* - 1)\Delta^{k+1},
\end{aligned}$$

where the last inequality follows that (55). Thus

$$\liminf_{\delta \to 1} h(b_i(w, \delta)) < \Delta^{k+1}. \tag{57}$$

But (55), (56), and (57) are mutually contradictory, so that Case 2 is impossible as well. ∎

The proof of the theorem can be completed with the use of a simple inductive argument. Fix δ^* as given by the previous lemma, consider any discount factor $\delta \in (\delta^*, 1)$, and a no-delay equilibrium at that discount factor.

First, consider all substructures $\pi \in \Pi^\circ$ such that $S(\pi)$ is a singleton. In this case, the subgame equilibrium must entail the formation of the singleton coalition. It is also clear that, for any RCF R, $R(\pi) = S(\pi)$ in this case, so that, in particular, equilibrium coalition formation is in agreement with some consistent RCF.

Inductively, suppose that the result is true for all substructures π such that $S(\pi)$ is of cardinality k or less for some $k \geq 1$. That is, on this subspace is defined an RCF R, the coalitions prescribed by which correspond to the equilibrium formation of coalitions. Pick some substructure $\pi \in \Pi^\circ$ such that $|S(\pi)| = k + 1$. At this stage, denote by x_i the equilibrium payoff to player $i \in S(\pi)$, were he to be the proposer.

Define a characteristic function $w_{R\pi}$ with player set $S(\pi)$, just as we did in (13):

$$w_{R\pi}(T) \equiv v(T; c(\pi \cdot T, R))$$

for all nonempty $T \subseteq S(\pi)$. Now follow the line of reasoning in the proof of Theorem 2.1. It is clear that, given the valuations $\{x_j\}$ defined in the previous paragraph, the maximum payoff that i can hope to achieve by making an acceptable offer is

$$x_i = \max_{T \subseteq S(\pi);\, i \in T} \left[w_{R\pi}(T) - \delta \sum_{j \in T \setminus \{i\}} x_j \right]. \tag{58}$$

By Lemma 5.6, this simply means that $x_i = a_i(w_{R\pi}, \delta)$ for all $i \in S(\pi)$. Moreover, i will make the offer to the coalition T_i in $\mathscr{C}_i(w_{R\pi}, \delta)$. By Lemma 5.7, T_i must lie in the set $\mathscr{C}_i(w_{R\pi})$ as well.

Now pick the first proposer assigned by the bargaining protocol to $S(\pi)$, say individual j. Pick the coalition T_j, and repeat this process for every substructure π such that $|S(\pi)| = k + 1$. This extends the RCF R to the set of all substructures π with $S(\pi)$ of cardinality at least $k + 1$, and completes the inductive step.

Once the induction is completed, we indeed have a consistent RCF that corresponds to equilibrium coalition formation at every stage for $R(\pi) \equiv T_j \in \mathscr{C}_j(w_{R\pi})$, as shown above. ■

APPENDIX: NONEXISTENCE OF A PURE-STRATEGY EQUILIBRIUM

The purpose of this section is to show that Theorem 2.1 cannot be strengthened to assert the existence of a pure-strategy equilibrium. To this end, we construct an example of the three-player game in which there is no pure-strategy equilibrium. The mixed-strategy equilibrium of this game will also serve to illustrate the notion of the mixed-strategy equilibrium used in Theorem 2.1. Finally, this example also makes the point that every equilibrium might involve a delay with positive probability.

Consider a three-player game in partition function form. We will denote by $v(\pi)$ the aggregate payoff to each of the coalitions in π. Thus $v(S, S') = (x, y)$ means that if the coalition structure is (S, S'), then coalition S gets an aggregate payoff of x and coalition S' gets y. Consider the specific description:

$$v(N) = 0 \qquad v(\{1\},\{2\},\{3\}) = (0.9, 0, 0) \qquad v(\{1,2\},\{3\}) = (0,0),$$
$$v(\{1\},\{2,3\}) = (0, 0.3) \qquad v(\{1,3\},\{2\}) = (1, 0.1).$$

PROPOSITION A.1. *Suppose* $1 > \delta > (8/9)^{1/3}$. *Consider any protocol such that if player* 2 *leaves the game, then the next proposer is player* 3 (*if still active*). *Then there is no pure-strategy, stationary equilibrium in this game. However, there does exist a mixed-strategy equilibrium* (*which also follows from Theorem* 2.1).

Proof. It will be useful to begin by making a couple of observations.

(1) Suppose there is an equilibrium in which x_i is the equilibrium payoff to i when i begins the game. Then, if in equilibrium, player i makes an acceptable proposal to a coalition containing player j, it follows that $x_j \geq x_i$; see Lemma 2 of Chatterjee *et al.* (1993).

(2) If, in equilibrium, player i makes an unacceptable proposal to player j, then this must result in player j's leaving the game with a coalition that does not contain i. Clearly, in equilibrium, player i will not make an unacceptable proposal to player j only to then accept a proposal form j. The claim then follows from the fact that, in equilibrium, there cannot be a chain of unacceptable proposals from i_1 to $i_2 \dots$ to i_k to i_1. Moreover, if i makes an unacceptable proposal to j who makes an unacceptable proposal to k, then $i \neq j$, k is better off saving one unit of time and making an unacceptable proposal to k rather than to j.

Let x_i be the expected equilibrium payoff to i if i starts as the first proposer. Clearly $0 \leq x_i \leq 1$ for all i. Since the protocol calls for 3 to make a proposal if 2 leaves, it follows that if player 2 leaves the game, then player 3 will offer 0.9δ to player 1, which will be accepted and, therefore, player 2 can obtain 0.1 by leaving the game. Thus $x_2 \geq 0.1$. However, it is easy to see that players 1 and 3 will get 0 if either one of them leaves the game unilaterally (if player 1 leaves the game, players 2 and 3 will form a two-person coalition).

CLAIM A.1. *As a proposer, player* 2 *will either leave the game or make an acceptable proposal to player* 3, *i.e.*, $x_2 = \max(0.1, 0.3 - \delta x_3)$.

Proof. As we have just observed, player 2 can receive 0.1 by leaving the game. If he makes an unacceptable proposal to player 1, by observation (2), it must be the case that player 1 will leave either alone or with player 3. In fact, player 1 will not leave alone since making an acceptable proposal to player 3 will yield $1 - \delta x_3 \geq 1 - x_3 \geq 0$, and one of these inequalities must be strict, whereas leaving alone will yield 0. But if 1 leaves with 3, player 2 gets 0.1δ. Thus player 2 will not make an unacceptable proposal to player 1. Making an unacceptable proposal to 3, by observation (2), will mean that player 2 gets 0 or 0.1δ, both which are dominated by 0.1. Clearly, leaving alone dominates making an acceptable proposal to player 1. The only other possibility is to make an acceptable offer to player 3 and receive $0.3 - \delta x_3$. Thus $x_2 = \max(0.1, 0.3 - \delta x_3)$ and player 2 will leave only if $\delta x_3 \geq 0.2$ and will make an acceptable offer to 3 only if $\delta x_3 \leq 0.2$.

CLAIM A.2. *As a proposer, player* 1 *will either make an acceptable proposal to player* 3 *or make an unacceptable proposal to player* 2, *i.e.*, $x_1 = \max(1 - \delta x_3, 0.9\delta^2\alpha)$, *where* α *is the probability with which player* 2 *leaves the game.*

Proof. If player 1 leaves the game, he gets 0 (because 2 and 3 will then form a two-person coalition). If he makes an unacceptable offer to player 3, by observation (2), either 3 leaves the game followed by 2, or 3 leaves with 2. Clearly, then, 3 must leave with 2, which will result in player 1 getting 0. If 1 makes an acceptable proposal to player 2, he cannot get more than 0. He can make an acceptable offer to player 3 and get $1 - \delta x_3 > 0$, which dominates all the other possibilities considered so far. The only remaining possibility is for him to make an unacceptable offer to player 2 and get $0.9\delta^2\alpha$.

Step 1. Suppose that $\alpha = 1$. It follows from Claim A.1 that $\delta x_3 \geq 0.2$ and $x_2 = 0.1$. Consider player 1's equilibrium strategy. Since $\delta \geq (8/9)^{1/3}$, we have $0.9\delta^2 > 0.8 \geq 1 - \delta x_3$. From Claim A.2, it now follows that player 1 will make an unacceptable offer to player 2 and $x_1 = 0.9\delta^2$. Now consider player 3's strategy. If he leaves the game, he gets 0. By observation (1), it cannot be the case that he makes an acceptable offer to player 2 (since $x_3 > 0.2 > 0.1$). Since player 1 makes an unacceptable offer to player 2, by observation (2), player 3 will not make an unacceptable offer to player 1. If he makes an unacceptable offer to player 2, he gets $\delta(1 - 0.9\delta)$. If he makes an acceptable offer to player 1, he gets $1 - 0.9\delta^3 > \delta(1 - 0.9\delta)$. Thus player 3's equilibrium strategy is to make an acceptable offer to player 1 and get $x_3 = 1 - 0.9\delta^3 < 0.2$. But this contradicts the presumption that $\delta x_3 \geq 0.2$.

Step 2. Suppose that $\alpha = 0$. Then, by Claim A.1, it follows that $\delta x_3 \leq 0.2$ and $x_2 = 0.3 - \delta x_3$. Since $\alpha = 0$, Claim A.2 implies that player 1 will make an acceptable offer to player 3 and get $1 - \delta x_3 > 0$. Now consider player 3's strategy. Since players 1 and 2 are making acceptable proposals to player 3, it follows from observation (2) that player 3 will either make an acceptable proposal to player 1 to receive $1/(1 + \delta)$ or make an acceptable proposal to player 2 and receive $0.3/(1 + \delta)$. Clearly, then, $x_3 = 1/(1 + \delta)$ and $\delta x_3 = 1/(1 + 1/\delta)$. Since $\delta > (8/9)^{1/3}$, we get $\delta x_3 > 1/((9/8)^{1/3} + 1) > 0.2$, a contradiction.

Step 3. From Steps 1 and 2 it follows that $\alpha \in (0, 1)$. This proves that there is no pure-strategy equilibrium in this game. Moreover a mixed-strategy equilibrium, with $0 < \alpha < 1$, much be such that $x_2 = 0.1$ and $x_3 = 0.2/\delta$. Consider player 1's strategy. By Claim A.2, $x_1 = \max(1 - \delta x_3, 0.9\delta^2\alpha) = \max(0.8, 0.9\delta^2\alpha)$. Thus $x_1 \geq 0.8 > x_3$ (since $\delta > (8/9)^{1/3}$). But now observation (1) implies that player 1 will not, in equilibrium, be making an offer to player 3. Thus player 1 will make an unacceptable offer to player 2 and $x_1 = 0.9\delta^2\alpha$. Finally, consider player 3's strategy. From the arguments in Step 2 it follows that player 3 must make an acceptable offer to player 1 and $x_3 = 1 - 0.9\delta^3\alpha$. Since $\delta x_3 = 0.2$, this yields

$$\alpha = \frac{\delta - 0.2}{0.9\delta^4}.$$

To summarize, then, a mixed-strategy equilibrium must be one in which $x_1 = 0.9\delta^2\alpha$, $x_2 = 0.1$, $x_3 = 0.2/\delta$, where player 2 leaves with probability $\alpha = (\delta - 0.2)/0.9\delta^4$ and makes an acceptable offer to player 3 with probability $1 - \alpha$. Player 1 always makes an unacceptable offer to player 2 and player 3 always makes an acceptable offer to player 1. It can be checked that this is, in fact, an equilibrium. ■

We shall end this section by showing that the conclusions derived from the above example remain valid even with a natural, weak form of superaddivity. When we consider a game in partition function form, our motivation comes from the supposition that such partition functions are "reduced versions" of game in strategic form. One of the most important restrictions that this imposes is the requirement of grand-coalition superadditivity. This means that the grand coalition should be able to achieve, in terms of aggregate worth, at least the sum of what is achievable under any coalition structure.

It turns out that one cannot impose, in general, any more than this final requirement on the grand coalition. That is, the superadditivity of *sub*-coalitions is *not* implied by games in strategic form.[23] To verify this

[23] This is not to say that further restrictions are not implied; only that it is not obvious what they are.

assertion, consider the Cournot example (Example 1.1) from the Introduction. The two-person coalition, confronted with a single opponent, can achieve strictly *less* than the sum of what two individuals can achieve in the three-person game.

So, in what follows, we will impose superadditivity at the level of the grand coalition but nowhere else. It should be emphasized, however, that our general results do not depend on making this assumption.

Consider a four-person version of this example, where the idea is that players 1 and 4 are interchangeable in the game. The partition function is as follows:

$$v(\{1\},\{2\},\{3\},\{4\}) = (0.9, 0, 0, 0.9),$$

$$v(\{42\},\{3\},\{1\}) = v(\{12\},\{3\},\{4\}) = (0, 0, 1),$$

$$v(\{13\},\{2\},\{4\}) = v(\{43\},\{2\},\{1\}) = (1, 0.1, 5),$$

$$v(\{1\},\{23\},\{4\}) = (0, 0.3, 0) \qquad v(\{1234\}) = 7,$$

and all other partitions have a zero vector of worths.

The protocol is as follows. Player 1 and 4 begins the game, but if there is any other player set S left with $3 \in S$, then player 3 is the first proposer in that set.

Notice that *if* player 1 (or 4) chooses to leave the game, then, according to the unique mixed-strategy equilibrium derived in the earlier version of the example, the remaining coalition structure will be $(\{34\}, \{2\})$ (or $(\{13\}, \{2\})$ if player 4 leaves), which yields the first leaving player a payoff of 5. On the other hand, if player 1 (or 4) does not leave and does *anything* else, then, in no equilibrium, can she get any return that is at least as high as 5. Therefore she will leave in equilibrium. This means that the equilibrium must involve some mixing (though the mixing is not observed on the equilibrium path).

REFERENCES

Aumann, R., and Myerson, R. (1988). "Endogenous Formation of Links between Players and of Coalitions: An Application of the Shapley Value," in *The Shapley Value: Essays in Honor of Lloyd Shapley* (A. Roth, Ed.), pp. 175–191. Cambridge: Cambridge Univ. Press.

Binmore, K. (1985). "Bargaining and Coalitions," in *Game Theoretic Models of Bargaining* (A. Roth, Ed.). Cambridge: Cambridge Univ. Press.

Bloch, F. (1996). "Sequential Formation of Coalitions with Fixed Payoff Division," *Games Econom. Behav.* **14**, 90–123.

Chakravorti, B., and Kahn, C. (1991). "Universal Coalition Proof Equilibrium," mimeo. Champaign, IL: Univ. of Illinois.

Chatterjee, K., Dutta, B., Ray, D., and Sengupta, K. (1989). "A Theory of Noncooperative Coalitional Bargaining," mimeo. College Station, PA: Pennsylvania State Univ.

Chatterjee, K., Dutta, B., Ray, D., and Sengupta, K. (1993). "A Noncooperative Theory of Coalitional Bargaining," *Rev. Econom. Stud.* **60**, 463–477.

Chwe, M. (1994). "Farsighted Coalitional Stability," *J. Econom. Theory* **63**, 299–325.

Dutta, B., Ray, D., Sengupta, K., and Vohra, R. (1989). "A Consistent Bargaining Set," *J. Econom. Theory* **49**, 93–112.

Dutta, B., van den Nouweland, A., and Tijs, T. (1995). "Link Formation in Cooperative Stituations," Discussion Paper, Tilburg Univ.

Greenberg, J. (1990). *The Theory of Social Situations*. Cambridge: Cambridge Univ. Press.

Hart, S., and Kurz, M. (1983). "Endogenous Formation of Coalitions," *Econometrica* **51**, 1047–1064.

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

Ichiishi, T. (1981). "A. Social Coalitional Equilibrium Existence Lemma," *Econometrica* **49**, 369–377.

Jackson, M., and Wolinsky, A. (1996). "A Strategic Model of Social and Economic Networks," *J. Econom. Theory* **71**, 44–74.

Krishna, V. and Serrano, R. (1996). "Multilateral Bargaining," *Rev. Econom. Stud.* **63**, 61–80.

Lucas, W. (1963). "On Solutions to n-Person Games in Partition Function Form," Ph.D. dissertation. Ann Arbor, MI: Univ. of Michigan.

Moldovanu, B. (1992). "Coalition-Proof Nash Equilibria and the Core in Three-Player Games," *Games Econom. Behav.* **4**, 565–581.

Okada, A. (1996). "A Noncooperative Coalitional Bargaining Game with Random Proposers," *Games Econom. Behav.* **16**, 97–108.

Perry, M., and Reny, P. (1994). "A Non-Cooperative View of Coalition Formation and the Core," *Econometrica* **62**, 795–817.

Ray, D. (1989). "Credible Coalitions and the Core," *Internat. J. Game Theory* **18**, 185–187.

Ray, D., and Vohra, R. (1996). "A Theory of Endogenous Coalition Structure," mimeo. Boston: Boston Univ.

Ray, D., and Vohra, R. (1997). "Equilibrium Binding Agreements," *J. Econom. Theory* **73**, 30–78.

Rosenthal, R. W. (1972). "Cooperative Games in Effectiveness Form," *J. Econom. Theory* **5**, 88–101.

Rubinstein, A. (1982). "Perfect Equilibrium in a Bargaining Model," *Econometrica* **50**, 97–109.

Salant, S., Switzer, S., and Reynolds, R. (1983). "Losses from Horizontal Mergers: The Effects of an Exogenous Change in Industry Structure on Cournot–Nash Equilibria," *Quart. J. Econom.* **93**, 185–199.

Selten, R. (1981). "A Noncooperative Model of Characteristic Function Bargaining," in *Essays in Game Theory and Mathematical Economics in Honour of Oscar Morgenstern* (V. Bohm and H. Nachtkamp, Eds.). Mannheim: Bibliographisches Institut.

Shenoy, P. (1979). "On Coalition Formation: A Game Theoretic Approach," *Internat. J. Game Theory* **8**, 133–164.

Thrall, R., and Lucas, W. (1963). "n-Person Games in Partition Function Form," *Naval Res. Logist. Quart.* **10**, 281–298.

Winter, E. (1993). "Mechanism Robustness in Multilateral Bargaining," *Theory and Decision* **40**, 131–147.

Zhao, J. (1992). "The Hybrid Solutions of an N-Person Game" *Games Econom. Behav.* **4**, 145–160.