Endogenous Formation of Coalitions in Noncooperative Games

José Luis Ferreira*

Department of Economics, Universidad Carlos III de Madrid 28903 Getafe, Madrid, Spain

In this paper we extend the definitions of coalition-proof Nash equilibrium and communication-proof equilibrium to situations of partial communication. Then we use the model of endogenous formation of coalitions of Aumann and Myerson to find that one of these situations in which the corresponding equilibrium exists will be endogenously determined by the players in the game. *Journal of Economic Literature* Classification Numbers: C70, C72

Key Words: Formation of coalitions, coalition-proof, communication-proof

1. INTRODUCTION

The definitions of subgame perfect equilibrium (SPE) (Selten, 1965), coalition-proof Nash equilibrium (CoalitionPNE) (Bernhein et al., 1987), and communication-proof equilibrium (Com-PE) (Ferreira, 1996) certainly reduce the number of equilibria in a given game with respect to the basic definition of Nash equilibrium (NE) (Nash, 1951); but they should not be regarded as refinements of NE. It is better to think of them as extensions of NE for different kinds of environments. In this way, SPE is an extension to games in extensive form, whereas CoalitionPNE and Com-PE are extensions for games with communication and no commitment for one stage and multistage games, respectively.

Here we present extensions (not refinements) of the CoalitionPNEa and Com-PEa that have been motivated by the two following discussions.

First, in noncooperative games, it has always been assumed that either no communication exists (NE and SPE), or that players can freely communicate but cannot sign agreements (CoalitionPNE and Com-PE). However, there are many intermediate possibilities between universal communication and no communication. This paper uses ideas from the literature of

cooperative games to interpret intermediate communication as free communication between players, but only within certain permitted coalitions. This is the first step toward a model of coalition formation. In cooperative games, where communication and the possibility of commitments are tacit assumptions, the formation of coalitions underlies several definitions like von-Neumann and Morgenstern's stable set (von-Neumann and Morgenstern, 1953). In applications of cooperative game theory, the topic is recurrent (see, for example, Shubik, 1982, and Mueller, 1989). However, nothing like this has appeared in the coalition and renegotiation proof literature.

Coalition-PNE is not the only definition that has been proposed to capture the possibility of coalitional deviations in normal form games, but it is, however, the first to incorporate some consistency criteria in the behavior of the intermediate and the grand coalitions. It is also the best known of all of the definitions, and it shares with the others the problem of nonexistence in many games. These are the main reasons to use this particular definition in the present paper, and the general analysis would not change if any other concept were to be used. For extensive form games, there are a couple of definitions that involve coalitional deviations, but only Com-PE includes renegotiation issues. Again, in the future other concepts might compete with Com-PE or would develop it further, but so far it is the only available definition, and, as in the normal form case, the main ideas of this paper do not depend on the particular definition.

Second, most definitions of equilibrium for games in extensive form require that the same kind of equilibrium be induced in every subgame, even in those that are not supposed to be played. This is a natural requirement if one wants to explain why the equilibrium path will be followed by all players. In the context of no communication, this requirement presents no problem, since existence has been proved for NE and most of its refinements under general conditions. For coalitional equilibria whose existence is not guaranteed, the requirement may be too demanding. The equilibrium may not exist in a continuation of the game that clearly will not be followed, but this situation should not be the cause for the nonexistence of the equilibrium in the whole game.

In this paper we use the model of endogenous formation of coalitions in Aumann and Myerson (1988) (A & M) to extend the definitions of CoalitionPNE and Com-PE to situations of partial communication, and provide existence of these extensions.

The general idea can be introduced with the following example. Suppose that three players play a game where $(10, 10, 0)$ are their respective payoffs in a self-enforcing coalitional equilibrium, when only players 1 and 2 can

communicate with each other before everyone chooses his or her strategy. Suppose further that $(8,8,8)$ are the payoffs of an equilibrium when the players of any two pairs of players (but not the other pair) can communicate with each other, and that no self-enforcing equilibrium exists when any coalition is possible. Finally, consider the situation in which players 1 and 2 meet first. One can argue that as long as the three players do not expect too much from a coalitional structure with no equilibrium (say no more than $(7,7,7)$), neither player 1 nor player 2 will open negotiations with player 3. This is not only because they will get 8 rather that 10 in that situation, but also because they know that once two pairs can communicate, the third pair will not open negotiations (each player would get 7 instead of 8). One may conclude that a structure with only two players being able to communicate and refusing further communications is a natural framework of negotiations for the game.

Following A & M, we will use the notion of a cooperation structure (also cooperation graph), introduced in Myerson (1977), to model the framework of negotiations for the game. This graph is defined as one whose vertices are the players. An edge of the graph is to be interpreted as a link between two players that, when it exists, can be used to carry on meaningful direct negotiations.

Modeling the endogenous formation of coalitions is equivalent to modeling the endogenous formation of links that form the cooperation structure. This can be done with a noncooperative game previous to the actual game, in which players decide with whom (if anyone) to open a link. Players must take into account that the decision of opening a link may lead other players to open new ones that were not previously profitable for them. Any equilibrium of this game provides a cooperation graph for the actual game. To anticipate the opening of new links and to decide whether that is good or bad, each player must know what to expect from every possible cooperation graph.

The differences between the model in A & M and the one presented in this paper start when a player compares his situation in two different graphs. In coalitional games, once a unique value is defined, the player has to compare just two numbers. In normal form games, we have no theory that predicts a unique coalitional self-enforcing equilibrium; moreover, the equilibrium may not exist.

The approach suggested in this paper is based on an extension to cooperation graphs of the definitions of CoalitionPNE and Com-PE, and on a definition of continuation values for the cases in which no equilibrium exists. These continuation values allow us to solve the auxiliary game in the same way as was discussed above, namely, by not insisting on having an

equilibrium behavior in every subgame. The role of these continuation values will be endogenously justified as a kind of fulfilled expectations.

The general conclusion of the paper is that for reasonable continuation values, the auxiliary game has a subgame perfect equilibrium that specifies a coalitional structure in which the CoalitionPNE (Com-PE) exists.

The model will be presented first for normal form games and then will be generalized to extensive form games. The generalization is straightforward, but is presented in a separate section because the extension to multistage games may not be the only natural one, and because the simplicity of the normal form makes the analysis simpler.

In recent years there has been a growing literature on coalition formation. A brief summary includes the study of noncooperative implementation of cooperative solutions (e.g., Pérez-Castrillo, 1994; Perry and Reny, 1994), alternative approaches to the model by Aumann and Myerson of coalition formation in cooperative games (e.g., Dutta, Nouweland and Tijs, 1996; Jackson and Wolinsky, 1996; and Ray and Vohra, 1996), and the introduction of hybrid models in which cooperative and noncooperative behaviors are allowed (e.g., Zhao, 1992). As far as we know, ours is the only totally non-cooperative study of coalition formation in which both the original and the auxiliary games are solved assuming a noncooperative behavior.

The paper is structured as follows. Section 2 presents the definitions, the model of coalition formation, and the results. In Section 3 some examples are analyzed. Section 4 contains a detailed discussion. Section 5 extends the analysis to multiperiod games, and Section 6 concludes.

2. A MODEL OF COMMUNICATION

Let g be a graph defined on N, the finite set of players. Given this cooperation graph g, let $C(g) = \{C \in 2^N \mid$ for all i and j in C, there exists a path in g restricted to C that goes from i to $j\}$; the elements in $C(g)$ will be interpreted as the coalitions that can be formed according to g. Now we can extend the CoalitionPNE to a cooperation graph g:

DEFINITION 1.   (i)   In a single player game $\Gamma = (N = \{1\}, S_1, u_1)$, a (in general mixed) strategy $s^* \in S_1$ is a Coalition-Proof Nash equilibrium restricted to a graph g, CoalitionPNE(g), if and only if s^* maximizes $u_1(s)$.

(ii)   Let $n > 1$, and assume that CoalitionPNE(g) has been defined for games with fewer than n players. Then

(a)   For any game $\Gamma = (N, S, u)$ with n players, $s^* \in S$ is self-enforcing restricted to g if, for all $C \in C(g)$, $C \neq N$, s_C^* is a CoalitionPNE(g) in the game $\Gamma|s_{-C}^*$ (where the strategies in s_{-C}^* are fixed).

4

(b) When $N \in C(g)$, for any game Γ with n players, $s^* \in S$ is a CoalitionPNE(g) if it is self-enforcing restricted to g and if there does not exist another $s \in S$ that is self-enforcing restricted to g and such that $u_i(s) > u_i(s^*)$ for all i in N.

The definition of CoalitionPNE(g) agrees with that of CoalitionPNE when g contains all possible edges. Then the Coalition-Proof Nash equilibrium may be interpreted as a self-enforcing equilibrium (if it exists) when there is full and free communication between players. Similarly, when g is the graph with no links, a CoalitionPNE(g) is a Nash equilibrium in Γ.

DEFINITION 2. For the game Γ, $e(g) = \{e_1(g), \ldots, e_n(g)\}$ is an aspiration profile in the cooperation graph g if

(i) $e(g) = \{e_1(g), \ldots, e_n(g)\} = \{u_1(s^*), \ldots, u_n(s^*)\}$ for some CoalitionPNE(g) s^*, if it exists.

(ii) For every $i \in \{1, \ldots, n\}$, $\min_s u_i(s) \le e_i(g) \le \max_s u_i(s)$ if no CoalitionPNE(g) exists.

In words, an aspiration for a player is his or her payoff in an equilibrium whenever it exists (all players expect the same equilibrium), and anything between the maximum and the minimum she or he can get in a situation of no equilibrium. In the latter case, players may have inconsistent beliefs. In Section 4 there is a discussion regarding the meaning of these aspirations.

The component $e_i(g)$ will be called an *aspiration* for player i in the graph g in the game Γ; and $e = \{e(g)\}_{g \in G}$, where G is the set of possible graphs on N, will be called a *set of aspirations* in the game Γ.

Given a normal form game Γ with n players, construct an auxiliary linking game as follows. Consider the set of pairs of players $N(N-1)/2$ and an ordering on this set (this ordering will be called the *rule of order*). Starting with the first player in the first pair, players decide whether to accept the opening of a link with the other player in the pair. To form a link, both potential partners must agree. Once formed, a link cannot be destroyed. The only other requirement is that after a link has been formed, each of the $n(n-1)/2$ pairs must be given an opportunity to form an additional link, starting with the pair next in line according to the rule of order. The game assumes perfect information. The game ends when, after opening a link, all of the pairs decide not to open a new one, when all links are formed or when no link is formed after the first round. This gives a standard finite extensive form game of perfect information, the result of which is a cooperation graph g. The payoff to each player i can be any aspiration $e_i(g)$.

Note that because the auxiliary linking game assumes perfect information, it has subgame perfect equilibria SPE in pure strategies for any set of

aspirations e. Each one of those equilibria is associated with a unique cooperation graph g, and therefore with a profile of aspirations in g, $e(g)$. To state the results, we need the following definition:

DEFINITION 3. Let g^* be the unique cooperation graph associated with a subgame perfect equilibrium in pure strategies of the auxiliary linking game of Γ when the aspirations are e^*. If $e^*(g^*)$ are the payoffs of a CoalitionPNE(g^*), then g^* is called a *natural structure* of Γ. The triple (g^*, s^*, e^*), where g^* is a natural structure, s^* is a CoalitionPNE(g^*) with payoffs $e^*(g^*)$, and e^* is a set of aspirations, will be called an *equilibrium in a natural structure* (ENS) for Γ.

For some given aspirations, a natural structure may not exist (when all subgame perfect equilibria of the auxiliary linking game lead to a cooperative graph g in which no CoalitionPNE(g) exists). Then the question is, what conditions on the set of aspirations e lead to natural structures in the auxiliary game? Proposition 1 provides an answer to this question. First define the sets $G^Y = \{g \in G \mid \text{CoalitionPNE}(g) \text{ exists}\}$ and $G_N = \{g \in G \mid \text{CoalitionPNE}(g) \text{ does not exist}\}$. Aspirations in a graph $g \in G_N$ will be called *continuation values*.

PROPOSITION 1. *Given the game Γ, there exists a profile of continuation values $\{e^*(g)\}$, $g \in G_N$, such that for any aspiration profile e satisfying $e(g) \leq e^*(g)$, there always exists a natural structure for Γ. Furthermore, these continuation values do not need to be lower than the payoffs of some natural structure existing in Γ for some other aspirations.*

Proof. First we establish the following fact. Given a game Γ, there exists an aspiration profile e^*, such that the auxiliary linking game has at least one subgame perfect equilibrium whose associated cooperation graph g^* is a natural structure of Γ.

This fact is proved by constructing one such profile of aspirations. For that, define an aspiration profile e with the condition that, for every cooperation graph $g \in G_N$, $e(g) = \{\min_s u_i(s)\}_{i \in N}$. Therefore, if we apply backward induction to find a SPE of the auxiliary game, at no point will a player find it profitable to choose an action that leads to a graph $g \in G_N$. In any SPE where players do not choose those actions, the resulting graph must belong to G_Y. The nonemptyness of G_Y is guaranteed by the existence of NEa.

Fact 2 is as follows. Let e be a profile of aspirations that lead to a natural structure g^*. Then any aspirations e' satisfying (i) $e'(g) = e(g)$ if $g \in G_Y$, and (ii) $e'(g) \leq e(g^*)$, otherwise, lead to g^*. To see this, consider the original auxiliary game with aspirations e, and the actions that constitute the SPE that leads to g^*. Consider these same actions when the aspirations are e'. If they do not form a SPE, an individual finds it

profitable to deviate at some point. If this deviation induces a graph in G_Y, then the same deviation would have upset the SPE in the original auxiliary game. If it induces a graph in G_N, the deviation is not profitable as, by definition, the player cannot improve his or her payoff with respect to g^*. The two contradictions complete the proof of fact 2.

Now notice that we can proceed as follows. Given a normal form game, define a first set of continuation values arbitrarily low for coalition structures in which no equilibrium exists, and find a first natural structure using the auxiliary linking game (by fact 1). Replace the first continuation values with the payoffs in this natural structure. Find a new natural structure with these new continuation values (by fact 2). We can continue in this way and get the existence of a natural structure with weaker conditions on the continuation values as long as we get increasing payoffs. This means that the continuation values do not need to be lower than the payoffs of some natural structure. Thus we have proved the proposition. Q.E.D.

A corollary of Proposition 1 is that for some aspirations not to lead to a natural structure, there has to be at least one player who expects more in some $g \in G_N$ than in some of the natural structures of Γ.

Since any CoalitionPNE(g) is a NE of the game, to get a natural structure it is enough that the continuation values (expectations in a nonequilibrium situation) are lower than the payoffs of a NE. This does not seem too demanding; however, there is no theory for expectations in nonequilibrium situations. To be on the safe side, one would like to be able to increase the upper bound for the continuation values in case the players have too optimistic (although incompatible) expectations in the case of nonequilibrium. The next proposition shows a way in which this can be done.

PROPOSITION 2. *Let g be a natural structure given by a SPE in the auxiliary linking game in which player i refuses to open a link with player j. Let g' be the graph resulting from adding the link (i, j) to the graph g. Then if $g' \in G_N$, the profiles $e_i(g') \leq e_i(g)$, $e_j(g') \leq e_j(g)$, $e_j(g')$ unbounded and $e_k(g') = e_k(g)$ ($k \neq i, j$) are continuation values that produce the same natural structure g.*

Proof. If i plays before j, it is a best action for j to choose to open a link with player i for large enough $e_j(g')$. Given this, the only way to prevent player i from choosing to open a link with player j is if $e_i(g') \leq e_i(g)$. Q.E.D.

There is still another way to relax the requirement for the continuation values. If we have that for a cooperation graph g there exists a CoalitionPNE(g) and that the same is true for any graph included in g, then we could use the payoffs in CoalitionPNE(g) as the upper bound of the continuation values to find a first natural structure. It would then be

important to know whether the existence of the solution concept that we apply to cooperation graphs has some monotonicity property of this kind with respect to graphs. Proposition 3 shows that, for at most three players, this property is indeed satisfied by CoalitionPNE.

PROPOSITION 3. *Let $N \leq 3$, and let g be a graph for which there exists a CoalitionPNE(g), then if $g^* \subset g$, there exists a CoalitionPNE(g^*). (Proposition 3 cannot be generalized for $N > 3$, as the following game shows:* $((A, B, C, D), \{S_i\}_i, \{u_i\}_i)$, with $S_i(i_1, i_2)$ $(i \in \{A, B, C, D\})$ and with $u(A_1, B_1, C_1, D_1) = (2, 2, 2, 2)$, $u(A_1, B_1, C_2, D_1) = (4, 4, 0, 0)$, $u(A_2, B_2, C_2, D_1) = (3, 3, 3, 0)$, $u(A_1, B_1, C_2, D_2) = (0, 0, 0, 1)$, $u(A_2, B_2, C_2, D_2) = (0, 0, 0, 1)$, and $u(A_i, B_i, C_i, D_i) = (0, 0, 0, 0)$ in all other cases.)*

Proof. If g^* is the graph with no links, then the CoalitionPNE(g^*) is the Nash equilibrium of the game.

Now let the set of players be (A, B, C), let g be the full graph, and let s^* be a CoalitionPNE(g).

(i) If there is only one link in g^*, say (A, B), and there does not exist a CoalitionPNE(g^*), then, in particular, s^* is not a CoalitionPNE(g^*). Then coalition (A, B) has a strategy (s_A, s_B) such that $s = (s_A, s_B, s_C^*)$ is preferred to s^* by both players A and B, and such that s is a CoalitionPNE in the game in which s_C^* is fixed; but then (s_A, s_B) is a deviation that contradicts the fact that s^* is a CoalitionPNE(g).

(ii) If there are two links in g^*, say (A, B) and (B, C), and if there exists a deviation by one pair, we repeat the argument before to conclude that s^* cannot be a CoalitionPNE. The only other possibility is a deviation by the three players, but then this deviation is a CoalitionPNE(g^*).

(iii) Finally, consider the case in which g is not the full graph: say (A, C) is not permitted, and g^* is a graph with one link, say (A, B). If s^* is the CoalitionPNE(g) and there is no CoalitionPNE(g^*), again, in particular, coalition (A, B) has a deviation as in (i) that is also a deviation from s.
                                                                    Q.E.D.

3. EXAMPLES

EXAMPLE 1 (Dividing a dollar by majority rule). In Bernhein et al. (1987), an example of nonexistence of CoalitionPNE is provided. In this section, this game will be solved by means of a natural structure.

Consider a three-player (A, B, C) game of dollar division, in which the allocation is decided by majority. That is, all players simultaneously an-

nounce allocations; if two or more players propose the same allocation, then that division is implemented, whereas if all disagree, the pie is discarded. Players wish to maximize the expected size of their shares. It is easy to check that the game has no CoalitionPNE, which in our terminology means that there is no CoalitionPNE(g) when g is the full graph in Fig. 1.

Define aspirations for this graph:

$$(1,1,1) \geq (e_A(g), e_B(g), e_C(g)) \geq (0,0,0).$$

Define the rule of order $\{(A,B), (B,C), (A,C)\}$. The graph g comes after all pairs have said yes to the formation of a link between their two players. Consider the situation when (A,C) is the last pair to decide whether to open a link or not. Player A has to compare $e_A(g)$ and his aspiration in the graph g_{AC} in Fig. 2. Similarly for player B. (Graphs g_{AB} and g_{BC} are defined in the same way.)

It is easy to see that A, B and C voting $(0,1,0)$ is a CoalitionPNE(g_{AC}), and that no other allocation can be achieved in any CoalitionPNE(g_{AC}). Therefore, after the graph g_{AC}, both A and C will be willing to open a link between them; in other words, g_{AC} leads to g.

Knowing this, let us take one more step backward. That is, knowing that (A,C) will open a link after g_{AC}, what will (B,C) do? When (B,C) have to decide after the graph g_C in Fig. 3, they have to keep in mind that any CoalitionPNE(g_C) gives payoffs of the form $(\lambda, 1-\lambda, 0)$, with $\lambda \in [0,1]$ (if player C gets a positive quantity, coalition (A,B) will deviate), and that after any two links, the third one comes immediately.

The complete analysis is shown in Fig. 4. The pairs of the form (i,j) in the figure mean that i chooses first (between Y and N), and then player j chooses knowing the choice of i. N is the continuation after one of them said N and Y is the continuation after the two of them said Y. For this particular example, the vectors $(\delta, 0, 1-\delta)$ and $(0, \gamma, 1-\gamma)$ are considered to be the respective payoffs of a CoalitionPNE(g_B) and of a Coali-
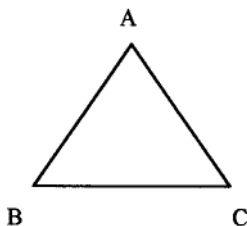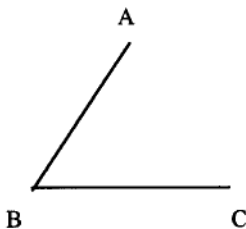

FIG. 1. Graph g.

FIG. 2. Graph g_{AC}.

tionPNE(g_A), where g_B and g_A are defined analogously to g_C. The vector (a, b, c) gives the payoffs of a Nash equilibrium: $(a, b, c) \geq (0, 0, 0)$ and $a + b + c = 1$. The arrows in the figure show the subgame perfect equilibrium for the case $\delta > \gamma$, $1 - \lambda > \gamma$, $\delta > a$, and $1 - \delta > c$. The other cases could be solved similarly. In any case it is assumed that $\lambda, \delta \geq e_A(g)$; $\lambda, 1 - \lambda \geq e_B(g)$; $1 - \gamma, 1 - \delta \geq e_C(g)$. The analysis predicts a natural structure of only one link joining two players who share the dollar and refuse to communicate with the third player.

By step 1 in the proof of Proposition 1, we can find a natural structure in the auxiliary game by setting very low continuation values, for instance, $(e_A(g), e_B(g), e_C(g)) = (0, 0, 0)$. With them we find the equilibrium depicted in the figure with payoffs $(\lambda, 1 - \lambda, 0)$, so that by Proposition 1 the continuation values may be as high as $(e_A(g), e_B(g), e_C(g)) = (\lambda, 1 - \lambda, 0)$. Since in the coalitional structure, player A refuses to open a link with player C, by Proposition 2 we can set $(e_A(g), e_B(g), e_C(g)) = (\lambda, 1 - \lambda, 1)$, which are incompatible expectations of what to get in the nonequilibrium situation.

The highest continuation values that we get by applying the propositions depend on, among other things, the rule of order. In the example, if players B and C were the first pair to meet, we could get the result $(0, \gamma, 1 - \gamma)$ for appropriate values of the parameters. Given this, the continuation values might be as high as $(1, \gamma, 1 - \gamma)$. Similarly, if players
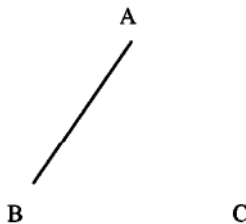

FIG. 3. Graph g_C.
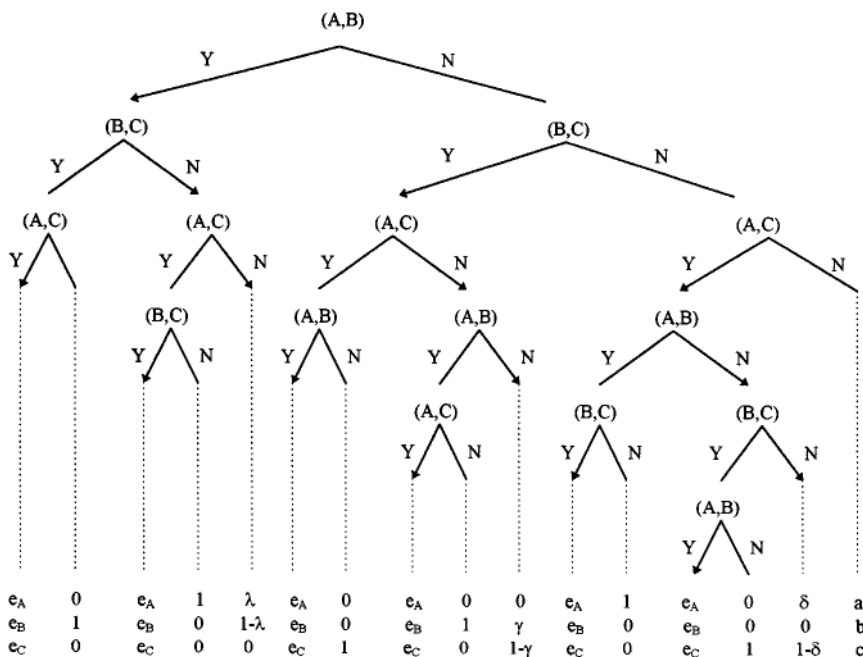
10

FIG. 4. Dividing a dollar by majority rule. Y, Both players in the pair agree to form a link. N, At least one player disagrees.

A and C met first, the continuation values could be $(\delta, 1, 1 - \delta)$. It might be sensible to require that the continuation values be the same in a given coalitional structure regardless of the rule of order. With this new requirement, we can set continuation values up to the levels (ν_1, ν_2, ν_3), where $\nu_1 = \min\{\lambda, 1, \delta\}$, $\nu_2 = \min\{1 - \lambda, \gamma, 1\}$, and $\nu_3 = \min\{1, 1 - \gamma, 1 - \delta\}$. Thus, if $\lambda = \gamma = \delta = 1/2$, $(\nu_1, \nu_2, \nu_3) = (1/2, 1/2, 1/2)$.

EXAMPLE 2 (One seller and two buyers). Consider the following two versions of the one seller (player A) and two buyers (players B and C) problem. In the first version, every player announces a division of one unit of outcome (worth one dollar). If the announcement of either B or C coincides with that of A, that division is implemented; otherwise, all three players get zero. In this game, every division may be achieved in a Nash equilibrium. If only one buyer and the seller can communicate, the coalition-proof Nash equilibria restricted to this graph have as payoffs a division of the dollar between the buyer and the seller. If both buyers can communicate with the seller, the seller gets everything. Whether the two buyers can communicate between them is irrelevant. One can construct

11

(A,B)

Y   N

(A,C)   (A,C)

Y N Y N

(A,B)

Y N

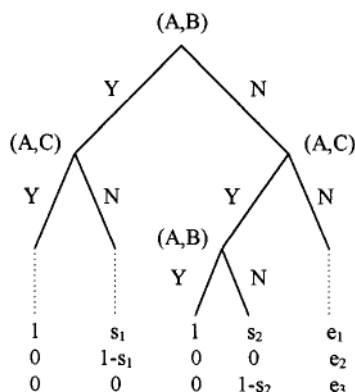| 1	s_1	1	s_2	e_1
0	$1-s_1$	0	0	e_2
0	0	0	$1-s_2$	e_3

FIGURE 5

the auxiliary linking game as in Example 1 to find the natural structure (with the same rule of order). Figure 5 shows this auxiliary game, in which the decisions to open a link between the two buyers have been omitted, as they do not add anything to the analysis.

There are three kinds of subgame perfect equilibria: (i) no links are open (both players in a last pair choose Y, and in the previous node they choose N), and a Nash equilibrium of the original game follows; (ii) only one buyer opens a link with the seller, and they share the dollar (if they are the last pair, both pairs choose N, and in the previous node they choose Y); and (iii) both buyers open a link with the seller and the seller gets everything (the pair (A,C) always chooses Y and the pair (A,B) first chooses Y and then N). The case studied in the example assumes $s_1 > s_2 > e_1$, $1 - s_1 > e_2$, and $1 - s_2 > e_3$; other cases are solved similarly. There are good interpretations for the first two equilibria, but not for the third. The first equilibrium can be obtained if, in the case of indifference, players make the decision to maximize the sum of payoffs of the pair in the link to be formed in that decision (maybe because ε-side payments are credible). The second situation is obtained if they decide not to open a link whenever they are indifferent (e.g., if to open a link is costly). To obtain that both buyers decide to open a link with the seller, it must be the case that, in her first decision node, player B behaves according to the first rule and, in his second node, according to the second rule, which is rather hard to interpret.

In the second game (second version), the two buyers make a bid, and the highest bid gets the dollar from the seller (in the tie-breaking rule, the winner is chosen at random). The only Nash equilibrium and the only

12

coalition-proof Nash equilibrium in this game is that both buyers bid one dollar, so that the seller gets all of the surplus. This being the case, anything can be a natural structure and the outcome is just that the seller gets one and the buyers get zero.

This example illustrates a case of two different games (institutions of trade) designed to solve the same economic problem that have the same outcomes in coalition-proof Nash equilibrium and that, nevertheless, have completely different solutions within the framework of natural structures.

4. DISCUSSION

4.1. *The Role of Continuation Values*

In the proposed model, the object that deserves more attention is the continuation values. Here we present several arguments in favor of their use.

(i) As we pointed out in the Introduction, one can construct a situation in which an equilibrium does not exist in subgames that clearly will not be followed. This should not be the cause for the nonexistence of the equilibrium in the whole game. An example is the game in which player one can choose between actions r and l. Action r gives a payoff of x_1, and action l leads to a subgame with no equilibrium. Consider the case in which her payoffs in the subgame are bounded by one. It is clear that for $x_1 > 1$, one will predict that player 1 will choose the outside option, even if the result of the subgame is unclear. However, according to most equilibrium concepts, the whole game has no equilibrium.

In different contexts there have been different ways to cope with nonexistence. In the example above, if the reason is that only pure strategies have been considered, the introduction of mixed strategies may be the solution. If the cause is, however, that the set of strategies is not compact, the use of ε-equilibria (where players maximize up to an ε) may restore existence. A third situation may be one in which the game is played with communication and a coalitional equilibrium is sought, but nevertheless the subgame does not have such an equilibrium. Again, even if x_1 is only a little smaller than 1, player one has to be too optimistic to engage in the subgame with no equilibrium. The introduction of continuation values allows us to evaluate how optimistic player one could be and still choose the outside option, so that the whole game has an equilibrium.

(ii) There are at least three other cases in game theory in which the out-of-equilibrium path is not clearly defined by the equilibrium concept. In the theory of social situations, by Greenberg (1990), a situation is regarded as (part of) an equilibrium if it is not dominated by another one.

13

In many instances this definition does not answer the question of what would actually happen out of the equilibrium. Again, when a game is solved using dominating strategies, the course of a subgame cannot be defined unless it is induced by these strategies. Finally, some set-valued solution concepts (like the stable sets of equilibria in Kohlberg and Mertens, 1986) may not be supported by one equilibrium continuation in every subgame, but by several continuations taken together as a set. In our case, we make explicit why an out-of-equilibrium path is not followed to understand the validity of the equilibrium.

(iii) When an equilibrium concept fails to exist in general, an immediate question is whether there exists a special set of games for which existence can be established. In our setting, the introduction of continuation values allows us to answer another interesting question: namely, what are the circumstances under which an equilibrium exists, not within the set of games, but within every game?

(iv) If one understands the resulting natural structure g and a CoalitionPNE(g) as an equilibrium for the game consisting of forming coalitions endogenously, and if one thinks that, at the end of that game, an equilibrium will be reached and that players also expect this to happen, then the continuation values are aspirations for an event that is not supposed to occur. Thus, sufficiently "nonequilibrium averse" aspirations (low enough continuation values) are consistent with the idea of expecting an equilibrium at the same time that they allow for the equilibrium to exist.

The situation is similar to that in sequential equilibrium, where any beliefs (and therefore many different expected payoffs) are allowed after a zero-probability event. In our case, we appeal to the idea that players either do not expect too much from a situation with no equilibrium, or do not expect a nonequilibrium at all.

Some sequential equilibria may be supported by extreme beliefs in zero-probability events. This has caused a literature on refinements of sequential equilibrium based on restricting the beliefs that may support an equilibrium, thus reducing its number. In our case, Propositions 1 and 2 show that imposing restrictions that play against existence (continuation values cannot be too low) and not imposing others that could play in favor (continuation values may be inconsistent) are still compatible with existence.

(v) Furthermore, like beliefs in sequential equilibrium, continuation values (and aspirations in general) are part of the equilibrium concept of natural structures. This favors the following interpretation: "not too high beliefs about payoffs from structure h support structure g, and therefore structure g is a possible equilibrium outcome of the game" instead of a different view that first predicts some beliefs and then a structure after them. (I thank a referee for this interpretation.)

4.2. *The Issue of Multiplicity and the Details of the Game*

The first example above showed a game with all players in a symmetric situation and with multiple equilibria compatible with this symmetry. If no coalitional deviations were possible, the outcome $(1/3, 1/3, 1/3)$ could be selected based on considerations of symmetry. However, when coalitions are possible, this outcome is simply unstable. The analysis of the game using natural structures translates the symmetry from the level of the payoffs (as in Nash equilibrium) to the level of the formation of winning coalitions. Now the problem is how to choose among the three different winning coalitions. The original game is silent in this respect. Either we are satisfied with this undetermined situation, or we use the theory to find out what kind of details that may be present in the reality can resolve it. The present paper favors the second approach as new details appear as important. These elements are exogenous to the original game and show a precise way (one precise rule of order and auxiliary linking game) in which players may endow themselves with a coalition structure in which the equilibrium exists, but the same conclusions could be obtained following other choices. Thus the arbitrariness of the exogenous elements in this work is not such a great problem.

Here we are not in a worse situation than the rest of the literature. Almost every refinement of Nash equilibrium is based on *ad hoc* assumptions that are not logically implied by the requirement of rationality. In the literature that is closest to our work, there is often an original game to be solved, but more structure is needed to give a hint of its resolution. For instance, in the case of noncooperative implementation of cooperative solutions, the non-cooperative game that provides the implementation is totally exogenous to the original one.

5. ENDOGENOUS FORMATION OF COALITIONS AND COMMUNICATION-PROOF EQUILIBRIUM

Definitions and propositions will be rewritten to accommodate multi-stage games. General extensive form games can also be analyzed if we consider that the number of stages is the maximum number of nested subgames in it; then a stage is the play between two consecutive nested subgames (see, for example, Bernhein et al. 1987). Let Γ be a multistage game (alternatively, an extensive form game) with a finite number of both players and stages. Define a cooperation graph g and the set of coalitions that are permitted in g, $C(g)$, as before. Let $\{n_1, \ldots, n_m\}$ be the set of initial nodes of subgames in Γ, and let g_i be the cooperation graph associated with node n_i. Then $E = \{g_1, \ldots, g_m\} \in \times_m G$ will be called a communication structure on Γ. Now we can extend the definition of Com-PE to a communication structure.

DEFINITION 1'. (i) In a single stage game Γ, s^* is a Com-PE(E) if and only if it is a CoalitionPNE(E). (Note that in this case, $E = \{g_1\}$, and Definition 1 applies.)

(ii) Inductively, let Γ be a game of t stages. Then s^* is a Com-PE(E) if and only if s^* is a CoalitionPNE(g_1) restricted to S' (Coalition-PNE(g_1, S')), where $S' = \{s \in S \mid s$ induces a Com-PE(E) in proper sub-games of $\Gamma\}$.

CoalitionPNE(g_1) restricted to S' simply means that all strategies (and deviations) must belong to the set S'. The definition of Com-PE(E) agrees with that of Com-PE when every g_i in E contains all possible edges, and with the definition of SPE when every g_i is the graph with no links. Aspirations are defined as in the one-stage case.

DEFINITION 2'. In the game Γ, $e(E) = \{e_1(E), \ldots, e_n(E)\}$ is an aspiration profile in the communication structure E if

(i) $\{e_1(E), \ldots, e_n(E)\} = \{u_1(s^*), \ldots, u_n(s^*)\}$ for some Com-PE(E) s^*, if it exists,

(ii) for every $i \in \{1, \ldots, n\}$, $e_i(E) \leq \max_s u_i(s)$ if no Com-PE(E) exists.

The component $e_i(E)$ will be called an aspiration for player i in the communication structure E, and $e = \{e(E)\}_{E \in \times G}$ will denote a set of aspirations for all possible coalition structures.

The auxiliary linking game for an extensive form game Γ is constructed as before, except for the additional rule that, at the beginning of each subgame, each pair has a new opportunity to open a link (following the rule of order).

DEFINITION 3'. Let E^* be the unique communication structure associated with a SPE in pure strategies of the auxiliary linking game of Γ when aspirations are e^*. If $e^*(E^*)$ are the payoffs of a Com-PE(E^*), then E^* is called a natural structure of Γ. The triple (E^*, s^*, e^*), where E^* is a natural structure and s^* is a Com-PE(E^*) with payoffs $e^*(E^*)$, will be called an ENS of Γ.

The counterparts of Propositions 1 and 2 can easily be stated and proved with these new definitions, but we can still improve the model. In Definition 3' we used the concept of SPE to find a communication structure. However, as links are open, some coalitions can be formed; therefore we could use the notion of Com-PE(E), where in each stage (subgame) the corresponding graph in E is determined by the pairs with a link at that point. Because the auxiliary linking game is of perfect information, the existence of a Com-PE(E) can be obtained as a corollary of the existence of Com-PE in this class of games (Ferreira, 1996).

EXAMPLE 3 (Dividing a dollar by majority with an outside option). Consider the example in which player 1 chooses between an outside option (action r) that gives him x_1, and a subgame g in which the game of dividing a dollar by majority rule as described above is played by him and players 2 and 3 (action l). The game in the example has no Com-PE, since there is no CoalitionPNE in the subgame g. However, different natural structures can be found. From the analysis above, we already know the natural structures for the subgame g that can arise under different specifications of the rule of order (Fig. 4). For simplicity, the equilibrium strategy in g will be identified with the vote (and shares) of the winning coalition, as the vote of the loser is irrelevant. The equilibria in natural structures for the complete game are as follows (for simplicity we omit the third element, the aspirations):

$$(\{C_1, (1,2)\}, \{r, (\lambda, 1 - \lambda, 0)\}) \qquad \text{if } x_1 \geq \lambda$$
$$(\{C_1, (1,2)\}, \{l, (\lambda, 1 - \lambda, 0)\}) \qquad \text{if } x_1 \leq \lambda$$
$$(\{C_2, (1,3)\}, \{r, (\delta, 0, 1 - \delta)\}) \qquad \text{if } x_1 \geq \delta$$
$$(\{C_2, (1,3)\}, \{l, (\delta, 0, 1 - \delta)\}) \qquad \text{if } x_1 \leq \delta$$
$$(\{C_3, (2,3)\}, \{r, (0, \gamma, 1 - \gamma)\}) \qquad \text{if } x_1 \geq 0$$
$$(\{C_3, (2,3)\}, \{l, (0, \gamma, 1 - \gamma)\}) \qquad \text{if } x_1 \leq 0.$$

The equilibrium in the first line is read as follows: the element C_1 is either $(1, 2)$ or the empty set, and is the communication structure in the first stage; the second element indicates the communication structure in the second stage (subgame g); the third is player one's action in his first move, and the last element is the (payoff of a) Com-PE in subgame g. The other equilibria are similarly interpreted, with $C_2 \in \{(1,3), \varnothing\}$ and $C_3 \in \{(2,3), \varnothing\}$.

EXAMPLE 4 (Divide a dollar by majority and then coordinate). This example consists of three players playing first the game of dividing a dollar and, in a second stage, the coordination game, in which $S_i = \{a_i, b_i\}$, $u_i(s) = 1$ if $s = (a_1, a_2, a_3)$, $u_i(s) = 2$ if $s = (b_1, b_2, b_3)$ and $u_i(s) = 0$ otherwise. The equilibria in natural structures are of the form

$$(\{(1,2), g_C\}, \{(\lambda, 1 - \lambda, 0), (2,2,2)\}),$$

where g_C is any graph at the beginning of the coordination game with the only condition that $(1, 2) \subset g_C$. The vector $(2, 2, 2)$ is the (payoff of the) Com-PE in the coordination game. The cases in which either pair $(1, 3)$ or pair $(2, 3)$ is formed at the beginning are similar.

EXAMPLE 5 (Coordinate and then divide the dollar by majority). Finally, if the coordination game is to be played before the game of dividing a dollar by majority, the equilibria in natural structures would be of the form

$$(\{\{(1,2),(1,2)\}, \{(2,2,2),(\lambda,1-\lambda,0)\}\}).$$

Now, no graphs with more than one link are formed, since that would lead to a situation of no equilibrium when the second game is to be played. Also note that $(1,1,1)$ cannot be the outcome in the first game; it certainly may be obtained as an equilibrium with the coalitional structure $(1,2)$, but players will form more links to deviate and get the outcome $(2,2,2)$, and that will result in a situation of no equilibrium in the second game. In other words, we have the interesting result that the presence of the second stage forces coordination among the three players, even if not all of them are linked in a communication structure.

6. CONCLUDING REMARKS

General existence of stable situations when coalitions can form is almost an impossibility, as the vote paradox, the core, and the CoalitionPNE show in different contexts. In this paper we showed that there is a way in which players can avoid this problem by endowing themselves with a cooperation structure in which the equilibrium exists. For every rule of order, this cooperation structure is endogenously formed.

The examples studied in the paper also illustrate the insight that natural structures may provide into very different issues, such as the relation between the symmetry of the game and the symmetry of the equilibrium (Example 1), the effects of natural structures on different mechanisms (games) designed to solve the same economic problem (Example 2), the coordination in efficient equilibria (Example 5), and the study and interpretation of social norms (Example 2).

These comments open two lines for future research. One is the study of different versions of the proposed model, among them the analysis of other types of auxiliary games and the imposition of more structure on beliefs in nonequilibrium situations. The other, and probably more interesting, is the finding of applications in which other issues (or the ones just mentioned) are addressed.

ACKNOWLEDGMENTS

the Tenth Italian Congress of Game Theory and Applications at Bergamo. The last version of this work was prepared while the author was visiting the MEDS Department at Northwestern University.

REFERENCES

Aumann, R. J., 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 S. Shapley*, pp. 175–191. Cambridge, England: Cambridge Univ. Press.

Bernhein, B. D., Peleg, B., and Whinston, M. D. (1987). "Coalition-Proof Nash Equilibria I: Concepts," *J. Econ. Theory* **42**, 1–12.

Dutta, B., Nouweland, A., and Tijs, S. (1996). "Link Formation in Cooperative Situations," mimeo.

Ferreira, J. L. (1996). "A Communication-Proof Equilibrium Concept," *J. Econ. Theory* **68**, 249–257.

Greenberg, J. (1990). *The Theory of Social Situations: An Alternative Game Theoretic Approach.* Cambridge, England: Cambridge Univ. Press.

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

Kohlberg, E., and Mertens, J. F. (1986). "On the Strategic Stability of Equilibria," *Econometrica* **54**, 313–327.

Mueller, D. (1989). *Public Choice II.* Cambridge, England: Cambridge Univ. Press.

Myerson, R. (1977). "Graphs and Cooperation in Games," *Math. Oper. Res.* **2**, 225–229.

Nash, J. (1951). "Noncooperative Games," *Ann. Math.*, **54**.

Pérez-Castrillo, D. (1994). "Cooperative Outcomes Through Noncooperative Games," *Games Econ. Behav.* **6**, 428–438.

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

Ray, D., and Vohra, R. (1996). "A Theory of Endogenous Coalition Structure," ISP Discussion Paper 68, Boston University.

Selten, R. (1965). "Spielthoretische Behandlung eines Oligopolmodells mit Nachfragetraheit," *Z. Ges. Staatswiss.* **121**.

Shubik, M. (1982). *Game Theory in the Social Sciences: Concepts and Solutions.* Cambridge, MA: MIT Press.

Von Neumann, J., and Morgenstern, O. (1953). *Theory of Games and Economic Behavior.* Princeton, NJ: Princeton Univ. Press.

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