

# Which Acceptable Agreements are Equilibria?\*

Sylvie Thoron

November 2002

First version December 2001

GREQAM,  
Centre de la vieille Charité,  
2, rue de la Charité,  
13002 Marseille,  
France.

Tel. : 0033491140741

fax number : 0033491900227

e-mail : thoron@ehess.cnrs-mrs.fr

---

\* The author would like to express her gratitude to Francis Bloch, Alessandra Casella, Jacques Durieux, Matthew Jackson, Alan Kirman, Phillippe Solal, Jörgen Weibull, to the participants in the PET Paris, SED New York, conferences, the members of IUI in Stockholm and an anonymous referee for helpful comments and stimulating discussions.

## **Abstract :**

To analyse the stability of a solution in normal form games of agreement formation, one has to specify what players believe will happen after a deviation. One possibility is to specify the conjectures that players may have in the rules of the game. The objective of this paper is to make the conjectures endogenous and consistent. To do this, I propose a normal form game in which each player's strategy is to say for each size of agreement whether it is acceptable or not. I propose a refinement, which guarantees that each one of these choices is self-enforcing. For general payoff functions, which exhibit positive externalities, I analyse situations in which symmetric players have the possibility to reach a unique agreement. I prove the uniqueness of this equilibrium. I give two specific examples: a cartel and an agreement to contribute to a public good.

*JEL Classification Numbers: C70, C72.*

*Key-words:* Coalition formation, normal form games, agreement, cartel, environmental agreement, public good.

## 1 Introduction

Consider a situation in which several parties have to decide whether or not to sign an agreement and that, if some of them reach an agreement, they will act in a way which will be beneficial for everybody, even for those who are not partners. Furthermore, suppose that this positive effect of the agreement increases with the number of partners. Obviously, in this kind of situation, before deciding on whether or not to participate, each party will pose a certain number of questions such as: How many people will sign with me? What happens if I do not sign? What should I announce as my intentions? What is acceptable for me, for the others?

The literature on stable cartels deals with some aspects of such situations. Indeed, the formation of a cartel in an oligopoly presents the kind of characteristics mentioned previously. For example, when a coalition of firms decide to form a cartel in a Cournot oligopoly, they decrease their production in order to increase the price. But the firms who remain independent benefit from this decrease of competition as well. D'Aspremont et al. (1983) made a first step towards a game theoretic analysis of this problem. In the case of symmetric firms forming a cartel, they propose a concept of stability. They say that a cartel is stable if the size of this cartel is internally stable in the sense that no partner has any incentive to leave it, and externally stable in the sense that no independent firm has any incentive to join it. They show that stable cartels always exist.

When d'Aspremont et al's model is reformulated as a game in which each player has a binary set of strategies and says yes or no to joining the cartel, a stable cartel is generated by a Nash equilibrium of this game. However, in their analysis, when a firm deviates, it does so on its own and it conjectures that the rest of the cartel will remain together and that no independent firm will react. In Thoron (1998), in order to take into account coalition deviations, I introduced a solution called the coalition proof stable cartel, using a concept of coalition proof Nash equilibrium (CPNE) proposed by Bernheim, Peleg and Whinston (1987). This refinement takes into account not only deviations by individuals but those by coalitions and furthermore, imposes a consistency

property on deviations. The idea is that a coalition deviation can only occur if it is credible, that is if there will be no further deviation by the original deviating members. I showed that the solution in the cartel game is unique. Note however, that this is only a step in the right direction since consistency is only required of the behaviour of deviating players. In contrast to those who deviate, the non-deviating players are assumed to be passive and not to modify their behaviour. Non-deviating partners stay in the cartel and non-deviating independent firms stay independent. Such behaviour corresponds to Nash conjectures. However, because of the positive externalities, there is a free riding problem and a partner whose conjecture is that the cartel will remain in place after her departure, may have an incentive to leave. In a Cournot game for example, the only Coalition Proof Stable Cartel is of size 1.

The role of conjectures in this framework is crucial. Any satisfactory analysis should go further and consider also the reaction of players other than those who initially deviate. To see this, we can compare the previous result with those obtained by Salant Switzer and Reynolds (1983) even though they use comparative static analysis rather than a game theoretic approach. In a symmetric oligopoly, they say that a merger is profitable if the merger's total payoff is larger than the sum of the members' payoffs before the merger, in the structure in which all the firms are independent. That is, they compare the situation with the existing coalition to one in which the coalition disappears. Using again the example of the Cournot oligopoly, they show that the profitable mergers are the largest ones, including the grand coalition. In the game theory framework, making their comparison, and looking for those cartels which are profitable in this sense, could be interpreted as looking for a stable cartel in which any member conjectures that her own deviation will provoke the complete collapse of the cartel. Thus, she has a specific conjecture about how the non-deviating members of the cartel will react to her deviation. The source of the difference between their results and those already mentioned (Thoron (1998)) is now clear. The fact that large coalitions are stable can be simply explained by the fact that this conjecture effectively dissuades any partner from deviating.

It should be clear now that the way in which we model games of agreement formation is still not fully satisfactory. What one would like is that conjectures about the reactions of other players should, in some sense, be rational. Up to this point only coalition deviations have been required to be rational or consistent. But one would also like these conjectures as to the reactions of all players to a change should be consistent.

Diamantoudi (2001) addressed this question in the framework of the theory of social situations. She modifies D'Aspremont et al (1983)'s internal and external stability concepts to propose a solution concept in the spirit of the abstract stable set. She envisages a sequence of reactions to an individual deviation. A cartel is internally stable if there does not exist a finite sequence of stable cartels such that a partner deviating from the initial cartel is, after the sequence of reactions, better off. External stability is revised in the same way. Applying this concept to a partition function derived from the original D'Aspremont et al. oligopoly in which the cartel is a price leader, she shows that there is a unique set of stable cartels.

In this paper, I want to deal with the problem of conjectures in the non-cooperative framework of strategic games. I propose a normal form game in which arbitrary conjectures about other players' reactions to deviations are no longer needed as a result of a richer specification of the strategy spaces. The idea is simple. Given the rules of the game, the strategy profile should allow one to determine what happens if a player or a group deviates.

To this end, I introduce a new notion of strategy, which is, for each player, to say, in the list of all the possible agreements, which ones are acceptable to her. This is rather natural since any player may find several different agreements acceptable. Indeed, in many situations in which a player is ready to negotiate an agreement, it is easy to imagine that she is willing to accept several possible agreements. Her objective is to reach the best choice available but since she does not know, a

priori, which agreement will form, she has an interest at the outset in specifying those which are acceptable. This is also the idea of the normal form game in which the players have to choose their strategies simultaneously.

Hence, for a given strategy profile, different sizes of agreements will be acceptable and feasible. However, it is assumed that the only feasible agreements which will be realised are those which are Pareto efficient. What happens if a player or group deviates from this strategy profile? She, or they, will do so by choosing a strategy which no longer results in the current agreements. Given that all the players have listed the agreements that they are prepared to accept, the new agreements which are now feasible are determined. Therefore, the new agreements which will be achieved depend on the new strategies of the deviating players, but also on the strategies of the non-deviating players. I introduce a specific refinement of the Nash equilibrium which ensures that the complete strategy of each player is self enforcing.

I specify rather weak general assumptions on payoff functions exhibiting positive externalities. Such functions may describe many examples other than cartels or mergers in which an agreement generates positive externalities. The same kind of phenomenon occurs, for example, when the agreement involves contributions to the production of a public good. That is why I will systematically use the generic term of agreement. Whatever the context, for any function satisfying these conditions, I can prove the uniqueness of the stable agreement and I give an algorithm to determine its size.

The properties of the general payoff function are presented in the following section. In Section 3, I present the game  $G$  of agreement formation and the results when Nash, Strong Nash and CPNE equilibrium are applied. Section 4 is devoted to the presentation of the strong Nash equilibrium in restricted games, the proof of its uniqueness and the algorithm which can be used for its characterisation. Section 5 clarifies the relationship between the game  $G$  and extensive form

games. Two examples with specific payoff functions, which satisfy the properties described in Section 2, are presented in section 6. Section 7, contains a discussion on a more general literature of coalition formation, on the assumptions made in the various approaches and on the possibility of extensions of the game  $G$ . I conclude in Section 8.

## 2 Payoff function With Positive Externalities

Consider a set of  $n$  symmetric players:  $N = \{1, \dots, n\}$ . They obtain a payoff which depends on how they are partitioned. In particular, I assume that the only partitions allowed are those with one coalition of  $k$  players and  $n-k$  independents. As a consequence of the symmetry of the players and this assumption of a unique agreement, the payoff of each player depends only on the size of the agreement and whether she is a partner to the agreement or independent. If she signs the agreement with  $k - 1$  partners,  $1 \leq k \leq n - 1$ , her payoff is denoted by  $\Pi^P(k)$ . On the other hand, if she is independent when  $k$  other players are partners in an agreement, her payoff is denoted by  $\Pi^I(k)$ .

The game is said to be essential: that is, there is at least one agreement of size  $k$ , which generates gains for its participants. Formally:

$$\exists k, 2 \leq k \leq n \mid \Pi^P(k) > \Pi^I(1).$$

Assume that the payoff function  $\Pi^I(k)$  satisfies two externality properties:

$$(P1) \quad \Pi^I(k) > \Pi^I(k - 1), \forall k = 2, \dots, n - 1.$$

$$(P2) \quad \Pi^I(1) = \Pi^P(1) \text{ and } \Pi^I(k) > \Pi^P(k), \forall k = 2, \dots, n - 1.$$

The first Property (P1) says that the payoff of an independent individual increases with the size of the agreement. This is the first effect of positive externalities. When the number of partners

increases, the agreement is more and more beneficial for those who remain independent. In fact, the agreement is beneficial for everybody. However it is also costly for the partners. Property (P2) means that whatever the size of the agreement, an independent individual always receives more than the partners to the agreement.

Now, consider the following definition:

**Definition 1** An agreement of size  $k$  is *advantageous* if the per member payoff is larger than the payoff in a society where all players are independent:

$$\Pi^P(k) \geq \Pi^I(1).$$

This definition corresponds to the notion of profitability introduced by Salant, Switzer and Reynolds [1983] in the framework of merger analysis. Here, I use it to define a third Property. Before I give the definition of this property, note that, because the game is essential, advantageous agreements exist and denote by  $k_1^*$  the smallest of these. Given this, I require the payoff function  $\Pi^P(k)$  to satisfy the following property:

$$(P3) \quad \Pi^P(k+1) > \Pi^P(k), \quad \forall k, k \geq k_1^*.$$

The payoff function  $\Pi^P(k)$  is not required to be increasing everywhere. However, Property (P3) guarantees increasing returns to collaboration from this minimum critical size  $k_1^*$ . Note that I do not impose any restriction on the second derivative of the payoff functions. These functions can be convex or concave or they may even have a non-monotonic second derivative, an *s*-shape function for example. I give two standard examples in Section 6, which satisfy these three Properties: a linear Cournot game and a public good contribution game.



### 3 The Game

In this Section, a normal form game of agreement formation denoted by  $G$  is presented. The set of  $n, n \geq 2$  players is  $N = \{1, \dots, n\}$ . They have to decide whether or not to sign an agreement. The outcome will be that at most one agreement will be signed. We saw in the previous section that, because the game is symmetric the players' payoffs can be defined as a function of the size of the agreement and of the position, partner or independent, of the player. Therefore, the number of partners in the agreement is the only relevant variable for each player's decision about her partnership. I assume that her strategy is to say which agreements she is willing to be a member of in terms of size. This means in particular that it is possible that a player is willing to sign agreements of different sizes. Formally, her strategy  $\sigma_i$  is an  $n-1$  vector in which the  $k$ th-component  $\sigma_{ik}$  is equal to one if  $i$  finds it *acceptable* to sign an agreement with  $k-1$  other partners or is equal to zero if she is not willing to do so. Therefore, the sets of strategies are:

$$\forall i \in N, \quad \Sigma_i = \left\{ \sigma_i = (\sigma_{i2}, \dots, \sigma_{in}) \mid \sigma_{ij} = 0 \text{ or } 1 \text{ for } j = 2, \dots, n \right\}$$

If, for a given player, her strategy is a  $n-1$  vector of zeroes, this means that this player is not willing to sign any agreement and that she wants to remain independent.

Each strategy profile  $\sigma$  is a matrix in which the row  $i$  is the strategy of player  $i$  and the component  $\sigma_{ij}$  is *the choice* of player  $i$  as to her partnership in the agreement of size  $j$ .

$$\sigma = \begin{bmatrix} \sigma_1 \\ \text{M} \\ \text{M} \\ \sigma_n \end{bmatrix} = \begin{bmatrix} \sigma_{12} & \mathbf{KK} & \sigma_{1n} \\ \text{M} & \mathbf{O} \quad \mathbf{N} & \text{M} \\ \text{M} & \mathbf{N} \quad \mathbf{O} & \text{M} \\ \sigma_{n2} & \mathbf{LL} & \sigma_{nn} \end{bmatrix}$$

$n \times (n-1)$

**Definition 2** Given a strategy profile  $\sigma$ ,

- i. An agreement of size  $k \geq 2$  is said to be *feasible* if:  $\sum_{i \in N} \sigma_{ik} \geq k$ .
- ii. It is *just feasible* if:  $\sum_{i \in N} \sigma_{ik} = k$ .

The meaning of this definition is that an agreement of size  $k$  is said to be *feasible* when at least  $k$  players find it acceptable and are willing to sign. Given this, the problem is now to know what is the outcome of the game? The first question that arises is to know what happens when the agreement of size  $k$  is feasible but not just feasible? In this case, more than  $k$  players would be willing to sign a  $k$  person agreement. I assume in this case, that the partners are picked at random among the players who find the agreement acceptable and that the payoffs are the expected payoffs before the members are chosen. Thus we have:

$$(1) \quad \tilde{\Pi}_i(k(\sigma)) = \begin{cases} \Pi^I(k(\sigma)) & \text{if } \sigma_{ik(\sigma)} = 0 \\ \rho(k(\sigma))\Pi^P(k(\sigma)) + (1 - \rho(k(\sigma)))\Pi^I(k(\sigma)) & \text{if } \sigma_{ik(\sigma)} = 1 \end{cases}$$

where  $\rho(k(\sigma))$  is  $i$ -player's probability of being a member of the agreement of size  $k(\sigma)$  when she finds it acceptable:  $\rho(k(\sigma)) = \frac{k(\sigma)}{\sum_{i=1}^n \sigma_{ik(\sigma)}}$ .

In order to define the outcome of the game, another question has to be answered. For a given strategy profile, *several sizes* of agreement can be feasible. Which one will be formed effectively? I assume a rule of efficiency. The Pareto efficient feasible agreements are the only ones, which have any chance of being signed, and if several agreements are efficient, each of them has the same probability. Given a strategy profile  $\sigma$ , denote by  $k(\sigma)$  a feasible agreement,  $F(\sigma)$  the set of feasible agreements and  $F^e(\sigma)$  the set of Pareto efficient feasible

agreements defined in the following definition. Denote by  $\pi(k)$  the structure composed by an agreement of size  $k$  and  $n - k$  independent players.

**Definition 3** For each efficient feasible agreement of size  $k' \in F^e(\sigma)$ ,

- i.  $\forall k'' \in F^e(\sigma)$ , it is impossible to increase one player's payoff without decreasing that of another, by changing from  $\pi(k')$  to  $\pi(k'')$ .
- ii.  $\forall k''$  such that  $k'' \in F(\sigma)$ , but  $k'' \notin F^e(\sigma)$ , it is always possible to increase at least one player's payoff without decreasing another player's payoff, by changing from  $\pi(k'')$  to  $\pi(k')$ .

**Definition 4 Efficiency Rule**

For a given strategy profile  $\sigma$ , all the elements of  $F^e(\sigma)$  are generated with equal probability.

We denote by  $\# F^e(\sigma)$  the cardinal of the set  $F^e(\sigma)$ . As a consequence of the efficiency rule and equation (1), we can define player  $i$ 's payoff for a given strategy profile  $\sigma$ :

$$(2) \quad \tilde{\Pi}_i(\sigma) = \frac{\sum_{k \in F^e(\sigma)} \tilde{\Pi}_i(k)}{\# F^e(\sigma)}.$$

Now that the normal form of the game is defined, I shall next look for the equilibria. As a first step, I characterise the set of stable agreements generated by the Nash equilibria of the game  $G$ , denoted by  $N$ .

**Definition 5**

- i. A strategy profile  $\sigma^*$  is a Nash equilibrium in the game  $G$  if and only if:

$$\tilde{\Pi}_i(\sigma_i^*, \sigma_{-i}^*) \geq \tilde{\Pi}_i(\sigma_i, \sigma_{-i}^*), \forall i \in N, \forall \sigma_i \in \Sigma_i$$

in which  $\sigma_{-i}$  is a vector giving each the strategy of each player except that of  $i$ .

- ii. An agreement of size  $k$  is stable if there exists a Nash equilibrium in the game  $G, \sigma^*$ , such that  $k \in F^e(\sigma^*)$ .

A first obvious necessary condition is given by the following lemma. It is a consequence of the fact that there is a minimum payoff that a player can guarantee for herself. Indeed, she can decide not to play the game and obtain in that case at least  $\Pi^I(1)$  since, according to property (P1), the payoff of an independent individual is an increasing function of the size of the agreement which is signed.

**Lemma 1** A stable agreement is always advantageous.

**Proof** First, note that, for a given strategy profile  $\sigma$ , each player's expected payoff is a weighted sum of payoffs  $\Pi^I(k)$  and/or  $\Pi^P(k)$  with  $k \in F^e(\sigma)$ . Therefore, her expected payoff increases if the player can eliminate the smallest of these payoffs.

Consider a strategy profile  $\sigma$ , and assume that  $\exists k_1, \dots, k_m$ , such that  $\forall t = 1, \dots, m, k_t \in F^e(\sigma)$  and  $k_t$  is not advantageous. This means that for each  $k_t$ ,  $\Pi^P(k_t) < \Pi^I(1)$ , and for every other size  $k \in F^e(\sigma), k \neq k_t, \Pi^P(k) \geq \Pi^I(1)$ . Now, each player who has chosen to be a partner in some of these non advantageous agreements can always increase her expected payoff by saying that these agreements are not acceptable. Indeed, by doing so, she eliminates the smallest possible payoffs  $\Pi^P(k_t)$  and increases the probability of larger payoffs:  $\Pi^I(k_t)$  or  $\Pi^P(k)$  and  $\Pi^I(k)$  when an advantageous agreement  $k$  is formed. Therefore,  $\sigma$  cannot be a Nash equilibrium.  $\square$

In spite of the fact that the efficiency rule permits a strategy profile to generate several agreements, the following lemma shows that in fact there will only be one agreement at the

equilibrium. For a given equilibrium strategy profile, there is only one Pareto efficient agreement, which is the largest feasible agreement when there are several of these.

**Lemma 2** At the equilibrium  $\sigma^*$ ,

- i. No more than one feasible agreement can be Pareto efficient:  $\# F^e(\sigma^*) \leq 1$ .
- ii. If several agreements are feasible, the Pareto efficient agreement is the largest one:

$$\text{If } k^* \in F^e(\sigma^*), \text{ then } k^* = \text{Max}\{k, k \in F(\sigma^*)\}.$$

**Proof**

i. Consider a given strategy profile  $\sigma$ . Assume that there is more than one Pareto efficient feasible agreement generated by  $\sigma$ ,  $\# F^e(\sigma) > 1$ . From lemma 1, all these agreements are advantageous. Consider the smallest of these agreements,  $\tilde{k} = \text{Min}\{k, k \in F^e(\sigma)\}$ .

As a consequence of property (P2):  $\Pi^P(\tilde{k}) < \Pi^I(\tilde{k})$ .

As a consequence of property (P3):  $\Pi^P(\tilde{k}) < \Pi^P(k') < \Pi^I(k'), \forall k' \in F^e(\sigma), k' \neq \tilde{k}$ .

For each player who finds  $\tilde{k}$  acceptable in  $\sigma$ ,  $\Pi^P(\tilde{k})$  is the smallest payoff she can obtain. Therefore, she has an incentive to deviate BY saying that  $\tilde{k}$  is not, in fact, acceptable and  $\sigma$  cannot therefore be an equilibrium.

ii. For a given equilibrium strategy profile  $\sigma^*$ , suppose that several agreements are feasible, that is,  $\# F(\sigma^*) > 1$ . Consider  $\hat{k} = \text{Max}\{k, k \in F(\sigma^*)\}$  the largest feasible agreement and assume that the unique Pareto efficient agreement is  $k^* \in F(\sigma^*), k^* < \hat{k}$ . This means that it is possible to increase at least one player's payoff without decreasing another player's payoff, by changing

from  $\pi(\hat{k})$  to  $\pi(k^*)$ . However, this is impossible since, as a consequence of (P1) and (P3), the independent individuals in  $\pi(\hat{k})$  prefer this structure to  $\pi(k^*)$ , whatever their position in  $\pi(k)$ :

$$\Pi^I(\hat{k}) > \Pi^I(k^*) > \Pi^P(k^*). \quad \mathbf{P}$$

The following proposition shows that the obvious necessary condition given in lemma 1 is also sufficient.

**Proposition 1** The set of stable agreements is the set of advantageous agreements.

$$N = \left\{ k \mid n \geq k \geq k_1^* \right\}$$

**Proof.** Consider an equilibrium strategy profile  $\sigma^*$ , which generates a set of just feasible agreements,  $F(\sigma^*) = \{k_1, k_2, \dots, k_{\hat{l}}\}$ ,  $k_1 < k_2 < \dots < k_{\hat{l}}$ . From lemma 2,  $F^e(\sigma^*) = \{k_{\hat{l}}\}$ . Assume that all the other agreements  $k$ ,  $k \neq k_t, t = 1, \dots, \hat{l}$  are unanimously considered as not acceptable:  $\sigma_{ik} = 0, \forall i \in N$ . Then, the only way a player  $i$  alone could improve her payoff, would be, when she finds  $k_{\hat{l}}$  acceptable, to provoke the collapse of this agreement and to become an independent. Since, then,  $F^e(\sigma_i, \sigma_{-i}^*) = \{k_{\hat{l}-1}\}$ , there is no incentive for player  $i$  to deviate if:

$$\Pi^P(k_{\hat{l}}) \geq \Pi^I(k_{\hat{l}-1}).$$

But for each advantageous agreement  $k_{\hat{l}}$ , it is always possible to find  $k_{\hat{l}-1}$  such that this inequality is verified, since it always holds for  $k_{\hat{l}-1} = 1$ . **P**

Note that, in the Example 2 presented in the previous section, which describes a contribution game to a public good, this implies in particular, that all the agreements are stable. At this point, a first question has to be answered. Is this result a consequence of the restriction to individual

deviations? With a Nash equilibrium concept, we cannot take into account deviations to form a larger agreement. However, in a game of coalition formation, it seems natural to use a concept taking into account coalition deviations. We have different choices among which the most famous are the Strong Nash Equilibrium and the Coalition Proof Nash Equilibrium (CPNE). The Strong Nash Equilibrium guarantees that no coalition can deviate, by changing the strategies of its members, when the strategies of the non-deviating players remain fixed.

**Definition 6**

(i) (Aumann (1959))  $\sigma^*$  is a Strong Nash Equilibrium in the game  $G$  if:

$$\begin{aligned} &\exists M \subset N \text{ and } \sigma_M \text{ such that :} \\ &\forall i \in M, \Pi_i(\sigma_M, \sigma_{-M}^*) > \Pi_i(\sigma_M^*, \sigma_{-M}^*). \end{aligned}$$

with  $\sigma_M$  the strategies of the  $M$ - members and  $\sigma_{-M}$  the strategies of the other players.

(ii)  $k^*$  is a strongly stable agreement if and only if there is a Strong Nash Equilibrium  $\sigma^*$  of the game  $G$  such that  $k^* \in F^e(\sigma^*)$ .

The following proposition gives the results when we apply this equilibrium concept to the game  $G$ . As far as notation is concerned, remember that  $k_1^*$ , defined in section 2, is the smallest advantageous agreement. Now, I introduce a new specific size. Let  $k_n$  be the first integer such that the inequality  $\Pi^I(k_n) > \Pi^P(n)$  is verified. Finally, let SN denote the set of strongly stable agreements

**Proposition 2** The set of strongly stable agreements of the game  $G$  is:

$$SN = \left\{ k \mid n \geq k \geq \text{Max}\{k_n, k_1^*\} \right\}.$$

**Proof.** Consider a Nash equilibrium  $\sigma^*$ , as defined in the Proof of Proposition 2, such that  $F^e(\sigma^*) = k_{\hat{I}}$ , with  $k_{\hat{I}} \geq \text{Max}\{k_n, k_1^*\}$ .

From (P3), when a player is partner in an advantageous agreement, her payoff increases with the size of the agreement. Therefore, no partner of the agreement  $k_{\hat{I}}$  wants to deviate to be in a smaller advantageous agreement. Furthermore, as a consequence of Property (P2), no independent player wants to be a partner in a smaller agreement. Formally:

$$\tilde{\Pi}_i(\sigma^*) = \begin{cases} \Pi^P(k_{\hat{I}}) \\ \text{or} \\ \Pi^I(k_{\hat{I}}) \end{cases} \quad \text{but} \quad \Pi^I(k_{\hat{I}}) \geq \Pi^P(k_{\hat{I}}) \geq \Pi^P(k), \quad \forall k \leq k_{\hat{I}}.$$

On the other hand, no independent player wants to deviate to be in a larger agreement. Indeed, by definition of  $k_n$  we know that:

$$\forall k_{\hat{I}} \geq k_n, \quad \forall k > k_{\hat{I}}, \quad \Pi^I(k_{\hat{I}}) \geq \Pi^I(k_n) > \Pi^P(n) \geq \Pi^P(k).$$

Therefore, each agreement of size  $k \geq \text{Max}\{k_n, k_1^*\}$  is strongly stable.

Now, consider the case in which  $k_n > k_1^*$ . Consider a Nash equilibrium  $\sigma^*$ , such that  $F^e(\sigma^*) = k_{\hat{I}}$ , with  $k_{\hat{I}} < k_n$ . Then, a coalition with all the partners in  $k_{\hat{I}}$  and all the independent player is ready to deviate from such an agreement since:

$$\forall k < k_n, \quad \Pi^P(k) \leq \Pi^I(k) \leq \Pi^I(k_n - 1) < \Pi^P(n).$$

Therefore, the agreements of size  $k \geq \text{Max}\{k_n, k_1^*\}$  are the only strongly stable agreements. **P**

**Definition 7** (Bernheim, Peleg, Whinston (1987))

$\sigma^*$  is a Coalition Proof Nash Equilibrium (CPNE), in the game  $G$  if and only if:

$$\sigma_i^* \in \text{argmax}_{\sigma_i \in \Sigma_i} \Pi_i(\sigma_i, \sigma_{-i}^*) \quad \forall i \in N$$



And for all  $M \subset N$  such that  $\# M = m$ , assuming that CPNE has been defined for games played by members of coalition  $J$  with fewer than  $m$  players, given  $\sigma_{-J}^*$ , then,

- $\sigma_M^*$  is self-enforcing; i.e.,  $\forall S \subset M$ ,  $\sigma_S^*$  is a CPNE in the game played by the members of  $S$ , given  $\sigma_{-S}^*$ ,
- there does not exist another self-enforcing strategy profile  $\sigma_M \in \Sigma_M$ , such that  $\forall i \in M$ ,  $\Pi_i(\sigma_M, \sigma_{-M}^*) > \Pi_i(\sigma_M^*, \sigma_{-M}^*)$ .

**Definition 8**  $k^*$  is a coalition proof stable agreement if and only if there is a CPNE  $\sigma^*$  of the game  $G$  such that  $F^e(\sigma^*) = \{k^*\}$ .

When they proposed this concept, Bernheim, Peleg, and Whinston (1987) wanted to take into account the fact that, in a non-cooperative framework, a deviation can only occur if it is credible, in the sense that it is itself immune against further deviations. I show in the following proposition that, taking into account this credibility does not change the results in the game  $G$ .

**Proposition 3** An agreement is coalition proof stable if and only if it is strongly stable.

$$CN = SN$$

**Proof.** We know that the Strong Nash equilibria are coalition proof (see Bernheim et al. (1987), p.7). Now, consider, when they exist, the agreements  $k \in [k_1^*, k_n]$ . I showed in the Proof of Proposition 2 that these agreements are not strongly stable because players have an incentive to deviate all together to form the grand coalition. Here, I can add that this deviation is credible in the sense of the CPNE, since the new agreement signed by the grand coalition is itself a CPNE. **p**

These last two results are rather strong. They mean that, in general, in the game  $G$ , taking into account coalition deviations is not enough to shrink the set of equilibria. We will see in the example given in Section 6, that it may happen that the three sets of Nash, Strong Nash and Coalition Proof Nash equilibria coincide exactly. Indeed, this is true as soon as  $k_n \leq k_1^*$ .

## 4 Restricted Games

In the game  $G$ , the problem, which arises when we apply the Nash equilibrium, is not only the problem of individual deviations. The characteristic features of this game are that, each player's strategy is to give a list of acceptable agreements and that, as we saw in the previous section, at the equilibrium, no more than one of these agreements will be formed. However, how can we interpret the other feasible agreements which do not have any direct effect on the final payoffs, and do we need to impose any requirements on them?

Consider the class of Nash equilibria characterised in the proof of Proposition 1. The idea is that when a player envisages leaving the agreement which is currently formed, her conjecture is that the agreement will disappear and therefore she does not move, simply to avoid finding herself in the situation in which all the players are independent. However, the player might reflect a little more and realise that since no player wants to wind up in this situation, the threat of the coalition breaking up is not credible.

Indeed, the other feasible agreements which are not formed at the equilibrium and have no direct effect on equilibrium payoffs, determine the conjectures of deviating players. Because we use a Nash equilibrium concept, when we consider a deviation, at the level of strategies, the conjecture is, of course, that the strategies of the others remain unchanged. However, even though other players do not change their strategies the coalition structure may still be modified. Players must then, have conjectures about how the structure would be modified after their move. I will require

that these conjectures be consistent. In order to deal with this problem, I introduce the notion of *restricted games*, defined on a set of strategies in which the maximum size of the coalition to which a player can choose to belong is limited.

When a game is restricted in the sense just described, the strategy space is modified. I will therefore need some new terminology to describe the components of the modified strategy spaces. Given  $\sigma_i$ , a player  $i$ 's strategy in the game  $G$ , a  $k$ -restricted strategy, denoted by  $\sigma_i^k$ , with  $k$  such that  $2 \leq k \leq n$ , is a restriction of the vector  $\sigma_i$  to its  $k - 1$  first components. It is also called a  $k$ -restriction of  $\sigma_i$ . Given a strategy profile  $\sigma$  in the game  $G$ , a  $k$ -restricted strategy profile denoted by  $\sigma^k$ , is a profile which associates to each player the  $k$ -restriction of her strategy. It is also called a  $k$ -restriction of  $\sigma$ .

$$\sigma^k = \begin{bmatrix} \sigma_1^k \\ \vdots \\ \sigma_n^k \end{bmatrix} = \begin{bmatrix} \sigma_{12} & \dots & \sigma_{1k} \\ \vdots & \ddots & \vdots \\ \sigma_{n2} & \dots & \sigma_{nk} \end{bmatrix}$$

$n \times (k - 1)$

**Definition 9** For each  $k, 2 \leq k \leq n$ , a *restricted game*  $G_k$  of the game  $G$  is a game with the same set of players  $N$  and the same payoff functions but in which, for each player  $i$ , the set of strategies is now:

$$\Sigma_i = \left\{ \sigma_i = (\sigma_{i2}, \dots, \sigma_{ik}) \mid \sigma_{ij} = 0 \text{ or } 1, j = 2, \dots, k \right\}$$

**Definition 10**

- (i) A strategy profile is an *Equilibrium in Restricted Games* (Nash, strong Nash or CPNE) of the game  $G$  if and only if  $\forall k, 2 \leq k \leq n$ , the  $k$ -restricted strategy profile is an Equilibrium in the restricted game  $G_k$ .

- (ii) An agreement  $k$  is stable in restricted games (stable, strongly stable or coalition proof stable) if there is an equilibrium (Nash, strong Nash or CPNE) in restricted games  $\sigma^*$  such that  $k \in F^e(\sigma^*)$ .

Let  $N^*$  be the set of stable agreements in restricted games. As we will see in the following proposition, taking into account individual deviations in restricted games is not sufficient to shrink the set of equilibria.

**Proposition 4** The set  $N^*$  of stable agreements in restricted games is equal to the set  $N$  of stable agreements which is the set of advantageous agreements:

$$N^* = N = \left\{ k \mid n \geq k \geq k_1^* \right\}$$

**Proof.** Consider an equilibrium strategy profile  $\sigma^*$ , which generates a set of just feasible agreements,  $F(\sigma^*) = \{k_1, k_2, \dots, k_{\hat{l}}\}$ ,  $k_1 < k_2 < \dots < k_{\hat{l}}$ . This is a Nash equilibrium in restricted games if and only if the condition defined in the proof of Proposition 1 is verified in every restricted game. For a given restricted game  $k$ , consider  $k_l$  the largest just feasible agreement not larger than  $k$ . The condition becomes  $\Pi^P(k_l) \geq \Pi^I(k_{l-1})$ . Therefore,  $\sigma^*$  is a Nash equilibrium in restricted games if:

$$(1) \quad \Pi^P(k_1) \geq \Pi^I(1) \text{ and } \Pi^P(k_l) \geq \Pi^I(k_{l-1}), \forall l = 2, \dots, \hat{l}.$$

However, if  $k_{\hat{l}}^*$  is advantageous, it is always possible to find a set of just feasible agreements  $F(\sigma^*) = \{k_1, k_2, \dots, k_{\hat{l}}\}$ , such that this is verified. The simplest example is  $F(\sigma^*) = \{k_{\hat{l}}\}$ .  $\blacksquare$

This first result might seem, at first sight, to be disappointing. However, as we will see in the next proposition, when restricted games and coalition deviations are combined, we can actually get uniqueness of equilibrium.

To do this I will first prove the following lemma, in the proof of which I characterise a specific sequence of critical sizes which I will use in Proposition 5. This is the sequence in which the first component is the smallest advantageous agreement and the other components are defined recursively. The size  $k_l$  is the smallest integer such that the payoff to a player who is in this agreement when it is formed is larger than an independent player's payoff when the agreement  $k_{l-1}$  is formed. In other words no individual in the agreement has an incentive to leave it and to become independent. This sequence is unique.

**Lemma 3** Under Properties (P1)-(P3), there is a unique sequence of critical sizes (integers)  $S^* = \{k_1^*, k_2^*, \dots, k_{\hat{l}}^*\}$ , in which the first component is the smallest advantageous agreement  $k_1^*$  and the other components are recursively defined as follows:

$$\forall l = 2, \dots, \hat{l}, \Pi^P(k_l^*) \geq \Pi^I(k_{l-1}^*) \text{ and } \Pi^P(k_l^* - 1) \leq \Pi^I(k_{l-1}^*).$$

The largest component of the sequence,  $k_{\hat{l}}^*$ , is such that:  $k_{\hat{l}+1}^* > n \geq k_{\hat{l}}^{*1}$ .

**Proof.** Because the game  $G$  is essential, there exists at least one advantageous agreement and  $k_1^*$  is the smallest one. Following Proposition (P2),  $\Pi^P(k_1^*) < \Pi^I(k_1^*)$ . However, (P3) says that  $\Pi^P(k+1) \geq \Pi^P(k), \forall k, k \geq k_1^*$ .

Now consider the following algorithm:

Denote by  $k_2^*$  the smallest integer  $k$  such that  $\Pi^P(k) \geq \Pi^I(k_1^*)$ . If  $k_2^* > n$  the sequence has only one component,  $k_1^*$ , otherwise it has at least two components,  $k_1^*$  and  $k_2^*$ . Then, continue to find  $k_3^*$  in the same way and so on. P

---

<sup>1</sup> When the payoff functions are concave, it can happen that  $k_{\hat{l}+1}^* \rightarrow \infty$ .

I give different examples in the following figures to illustrate Lemma 3:

Figures 1,2,3

Denote by  $SN^*$  (respectively  $CN^*$ ) the set of strongly (coalition proof) stable agreements. The result of the following proposition is that, when we combine restricted games and coalition deviations, we obtain uniqueness of the equilibrium. Indeed, the agreement of size  $k_{\hat{l}}^*$  is then the unique strongly stable agreement in restricted games, and the unique coalition proof stable agreement in restricted games.

**Proposition 5** There is a unique strongly (and coalition proof) stable agreement in restricted games, which is the largest element of the sequence  $S^*$ .

$$SN^* = CN^* = \{k_{\hat{l}}\}$$

**Proof.** i. Consider an equilibrium strategy profile  $\sigma^*$ , which generates a set of just feasible agreements,  $F(\sigma^*) = \{k_1, k_2, \dots, k_{\hat{l}}\}$ ,  $k_1 < k_2 < \dots < k_{\hat{l}}$ . As we saw in the proof of Proposition ,  $\sigma^*$  is a Nash equilibrium in restricted games if and only if:

$$(1) \quad \Pi^P(k_1) \geq \Pi^I(1) \text{ and } \Pi^P(k_l) \geq \Pi^I(k_{l-1}), \forall l = 2, \dots, \hat{l}.$$

In the proof of Proposition 2, I showed that, from a Nash equilibrium, the only possible group deviation is the joint deviation of independent players and the partners in  $k_{\hat{l}}$  to form a larger agreement. The condition  $k_{\hat{l}} \geq k_n$ , where  $k_n$  is the first integer such that the inequality  $\Pi^I(k_n) \geq \Pi^P(n)$  is verified, guarantees that this cannot occur in the game  $G$ . Indeed, we have in that case  $\Pi^I(k_{\hat{l}}) \geq \Pi^I(k_n) \geq \Pi^P(n)$  and the independent players do not want to deviate.

For each given restricted game  $G_k$ , where  $k_{\hat{l}} \leq k < n$ , it is still true that no deviation is possible since an independent player could have at most  $\Pi^P(k) < \Pi^P(n)$ .

For each given restricted game  $G_k$ , with  $k_{\hat{l}-1} \leq k < k_{\hat{l}}$ , the condition becomes:  $k_{\hat{l}-1} \geq k^*$ , with  $k^*$  the first integer such that the inequality  $\Pi^I(k^*) \geq \Pi^P(k)$  is verified. Indeed, in that case we have  $\Pi^I(k_{\hat{l}-1}) \geq \Pi^I(k^*) \geq \Pi^P(k)$ , and no group deviation is possible in the restricted game. From (P3) we know that this condition is verified in each restricted game  $G_k$ , with  $k_{\hat{l}-1} \leq k < k_{\hat{l}}$  iff:

$$(2) \quad \Pi^I(k_{\hat{l}-1}) \geq \Pi^P(k_{\hat{l}} - 1).$$

Therefore, we want the set of feasible agreements  $F(\sigma^*) = \{k_1, k_2, \dots, k_{\hat{l}}\}$  to verify both (1) and (2) conditions in each restricted game. As a consequence  $k_1 = k_1^*$ , and:

$$(3) \quad \forall l = 2, \dots, \hat{l}, \Pi^P(k_l) \geq \Pi^I(k_{l-1}) \text{ and } \Pi^P(k_l - 1) \leq \Pi^I(k_{l-1}).$$

However, there is a unique set of feasible agreements that verifies these inequalities, this is the sequence  $S^* = \{k_1^*, k_2^*, \dots, k_{\hat{l}}^*\}$  defined in lemma 3.

ii. Note that, if the condition (2) is not satisfied, the deviation can occur and it is credible since no further deviation would be beneficial. ■

Therefore, in the normal form game  $G$ , the uniqueness of the equilibrium agreement is obtained using a specific refinement, the strong Nash equilibrium in restricted games. At this point, natural questions would be: What is the relationship between this specific equilibrium refinement and subgame perfection in an extensive form game? Does an extensive form game of agreement formation exist, which coincides with the normal form game  $G$ ?

## Section 5 Relationship with Extensive Form Games

Indeed, as will be proved in the following proposition, this extensive form game exists, and is specified as follows. Call this game  $\Delta_1$ . The players are now ordered following a protocol  $\rho$ . The

first player, player 1 in the protocol  $\rho$ , starts the game and chooses an integer  $k \in [1, \dots, n]$ . If  $k = 1$ , the player leaves the game and the player 2 has to choose an integer  $k \in [1, \dots, n - 1]$ . If the player chooses an integer  $k > 1$ , a coalition of size  $k$  is formed with this player and the  $k - 1$  following players in the protocol. The game stops as soon as one player has chosen an integer larger than 1 or, if this does not happen, when all the players have made the choice  $k = 1$ . Therefore, for player number  $m$  in the protocol, her strategy is the choice of an integer  $k \in [1, \dots, n - m + 1]$ . The outcome is a structure with at most one coalition of size  $k$  and  $n - k$  independent players.

The game  $\Delta_1$  is a version of the game  $\Delta$  proposed by Bloch (1996): a bargaining game of coalition formation between symmetrical players. His Proposition 4.2 p. 109 shows that, because of the symmetry of the players, this game is equivalent to a more general bargaining game in which the players may be heterogeneous. In this game, the players in the order determined by the protocol make a proposition as to the coalition they wish to join. If the other members of a proposed coalition accept, that coalition forms and leaves the game. The procedure is then repeated with the remaining players. On the other hand, in the game  $\Delta$ , as soon as a player proposes an integer  $k$ , she can form a coalition with the  $k - 1$  following players (in the protocol) because they automatically accept. The difference between the game  $\Delta_1$  and Bloch's game  $\Delta$  is that, in the game  $\Delta_1$ , the game stops when at most one coalition has been formed. I prove the following proposition, which gives the relationship between the results of the extensive form game  $\Delta_1$  and the normal form game  $G$  presented before.

**Proposition 6** The sub-game perfect equilibrium in pure strategies of the game  $\Delta_1$  coincides with the strong Nash equilibrium in restricted games of the game  $G$ . Both generate the agreement  $k_l^*$ .



**Proof.** I characterise the sub-game perfect equilibrium of the game  $\Delta_1$  by backward induction. The only possible strategy for the last player  $n$  is to remain alone. However, as long as  $k$  is not advantageous, i.e. as long as  $k < k_1^*$ , player number  $n - k + 1$  in the protocol, prefers also to remain in the structure in which everybody is independent. She stays alone and knows that the following players will do the same. Then, player  $n - k_1^* + 1$  chooses  $k_1^*$  since it is better to be a partner in this agreement than to be in the trivial structure. For any  $k$  larger than  $k_1^*$  where  $\Pi^I(k_1^*) \geq \Pi^P(k)$ , player  $n - k + 1$  chooses 1, and waits for the formation of  $k_1^*$ . Then, player  $n - k_2^* + 1$  chooses  $k_2^*$ , which is the first integer such that  $\Pi^I(k_1^*) < \Pi^P(k_2^*)$ . Proceeding in this way, we characterise the sequence  $S^*$  defined in lemma 3 and the only coalition which is formed is  $k_1^*$ , the first one to be proposed. ■

Of course, it is always possible to express any given game, either in extensive form or in normal form. However, although their equilibria coincide, the games  $\Delta_1$  and  $G$  are not equivalent. Usually, when we want to write the normal form of an extensive form game, we have to specify each player's strategy giving an action for each subgame in which she has to move. In the game  $\Delta_1$ , each player moves in only one subgame. Remember that the game  $\Delta_1$  is characterised by a protocol. Denote by  $\Delta_1^p$  the game characterised by the protocol  $p$ . When the player number  $m$  in the protocol  $p$  moves, this is because the  $m - 1$  first players in the protocol have chosen to remain independent and there are  $n - m + 1$  players left in the game. This characterises the sub-game. In the game  $G$  there is no protocol, but we can view each player's strategy as the choice of an action for each protocol of the game  $\Delta_1$ . When a player has to say if she finds acceptable to sign an agreement of size  $k$ , this is as if  $n - k$  players have left the game.

To conclude this section, it is worth making the following remark. It might be tempting to consider the sequence  $S^*$  as corresponding to the sub-game perfect Nash equilibrium of an

extensive form corresponding to  $G$ . To pursue this analogy, one would consider the restricted games  $G_k$  in increasing order which would consider with the idea of backward induction. However, doing this does not yield the sequence  $S^*$  and indeed, any advantageous agreement becomes stable. What restricts the equilibria of the game  $G$  is not the order in which the restricted games are considered but rather the introduction of coalition deviations and thus of a strong Nash equilibria concept. Hence, a sub-game perfect *Nash equilibrium* of the bargaining game  $\Delta_1$  coincides with a *strong Nash equilibrium* in restricted games of the normal form game  $G$ .

## Section 6 Applications

In this section, I give two specific examples that yield payoff functions satisfying Properties (P1)-(P3) and which are commonly used in the literature.

### *Example 1: Cournot game and cartel formation*

Consider the simplest oligopoly game “à la Cournot” with zero cost, a homogeneous good, and a linear demand function  $P = 1 - Q$ , where  $P$  is the price and  $Q$  is the total quantity produced. Assume that there are  $n$  firms and that  $k$  of them form a cartel. As a result of the competition between the cartel of  $k$  firms and the  $n - k$  firms which remain alone, an independent firm's profit is

$$\Pi^I(k) = \frac{1}{(n - k + 2)^2},$$

and a cartel member's profit is

$$\Pi^P(k) = \frac{1}{k(n - k + 2)^2}.$$

Properties (P1) and (P2) are obviously verified. Furthermore,  $\Pi^P(k)$  is a convex function and property (P3) is verified. Now compute  $k_1^*$  as a root of the following equality:

$$\Pi^P(k) = \Pi^I(1) \Leftrightarrow k(n - k + 2)^2 = (n + 1)^2.$$

We obtain:

$$k_1^* = \frac{2n+3 - \sqrt{4n+5}}{2}.$$

By definition, we could find  $k_2^*$  by computing the root of the following equality:

$$\Pi^P(k) = \Pi^I(k_1^*).$$

However, it is easy to verify that:  $\Pi^I(k_1^*) \geq \Pi^P(n) = \frac{1}{4n}$ . We deduce that  $k_2^* > n$  since  $\Pi^P(k)$  is an increasing function for  $k$ ,  $k \geq k_1^*$ .

The other consequence is that  $k_n$  defined such that  $\Pi^P(n) = \Pi^I(k_n)$  is smaller than  $k_1^*$ .

Therefore, all the advantageous agreements of size  $k \geq k_1^*$  are strong Nash equilibria, whereas the

unique SNERG is  $k_1^* = \frac{2n+3 - \sqrt{4n+5}}{2}$ .

Note that, in this case, the set of agreements sustained by a strong Nash equilibrium corresponds to the set of profitable cartels, following the definition proposed by Salant, Switzer and Reynolds (1983).

Note also that, the SNERG corresponds to the equilibrium of the bargaining game  $\Delta$  of coalition formation proposed by Bloch (1996). When this game  $\Delta$  is applied to Cournot payoff functions, only one coalition is formed. Then, as consequence of Proposition 6, this is the same coalition as the coalition obtained with the game  $G$  (and hence  $\Delta_1$ ).

### Example 2: *Contribution to a public good*

In this example, I consider a classic model of public good agreement. This model has been presented by Yi (1997) to illustrate cases in which there are positive externalities. Ray and Vohra (2001) use this payoff function in their bargaining game of coalition formation. Each player is endowed with 1 unit of a private good. At cost  $c(x_i)$ , agent  $i$  can provide  $x_i$  units of the public good. Let  $X = \sum_{i=1, \dots, n} x_i$  be the total amount of the public good. Each player enjoys the same benefit from consuming the public good,  $g(X)$ . Player  $i$ 's net utility is given by  $g(X) - c(x_i)$ .

Assume that:  $g(X) = X$  and  $c(x) = cx^2$ ,  $c > 0$ . Suppose that  $k$  players sign an agreement denoted by  $K$ . As a consequence of this agreement, the members choose their provision of the public good to maximise their joint utility:  $kg(X) - \sum_{j \in K} c(x_j)$ . Simultaneously, each of the  $n - k$  non-members chooses her provision of the public good to maximise her own net utility  $g(X) - c(x_i)$ .

At the equilibrium, for a given  $k$ , a member's payoff is:

$$\Pi^P(k) = \frac{1}{4c} \{2(n - k + k^2) - k^2\}$$

and a non-member's payoff is:

$$\Pi^I(k) = \frac{1}{4c} \{2(n - k + k^2) - 1\}.$$

Obviously, Properties (P1), (P2), (P3) are verified. The two payoff functions are increasing and for a given  $k, 1 \leq k \leq n$ ,  $\Pi^I(k) \geq \Pi^P(k)$ .

We can find the strong Nash equilibria by solving the following equality:

$$\Pi^I(k_n) = \Pi^P(n).$$

This amounts to solving the following polynomial:

$$2k_n^2 - 2k_n - (n-1)^2 = 0.$$

The positive root is:

$$(R_1) \quad k_n = \frac{1 + \sqrt{1 + 2(n-1)^2}}{2}.$$

The set of agreements associated with the strong Nash equilibria is the set of agreements larger or equal to  $k_n$  given by (R1).

We can find the different components of the sequence  $S^* = \{k_1^*, k_2^*, \dots, k_l^*\}$  by solving the following equation:

$$\Pi^P(k_l) = \Pi^I(k_{l-1}) \Leftrightarrow 2(n - k_l + k_l^2) - k_l^2 = 2(n - k_{l-1} + k_{l-1}^2) - 1.$$

This amounts to solving the following polynomial:

$$k_l^2 - 2k_l - F(k_{l-1}) = 0 \quad \text{with} \quad F(k_{l-1}) = 2k_{l-1}(k_{l-1} - 1) - 1$$

One of the two roots is always negative. Therefore, the solution is:

$$(R_2) \quad k_l = 1 + \sqrt{2k_{l-1}(k_{l-1} - 1)} = f(k_{l-1}).$$

In fact, of course, the elements of the sequence are integers. The smallest element is  $k_1^*$ , the smallest advantageous agreement. Let  $e[x]$  be the largest integer smaller than  $x$ . For each  $l$ ,  $1 \leq l \leq \hat{l}$ , the  $l$ th element of the sequence is then:

$$(R_3) \quad k_l^* = e\left[f\left(k_{l-1}^*\right)\right] + 1.$$

Note that the slope of the payoff functions does not depend on  $n$  and, as a consequence, the different elements do not depend on  $n$  either. However, the size of the agreement associated with a strong Nash equilibrium in restricted games,  $k_{\hat{l}}^*$ , depends on  $n$ . Indeed, this is the largest element of the sequence smaller than  $n$ . I give the values of the first components in the following table.

**Table 1:** Strong Nash equilibrium in restricted games

This table should be read as follows: the second row gives the value of the corresponding element of the sequence. The third row gives the values of  $n$  such that the corresponding element is the size of the stable agreement.

Here, we can make the following remark. For  $n < 9$ , this sequence corresponds to the special collection of integers characterised by Ray and Vohra (2001). The equilibrium coalition structure of their game is then what they call the *decomposition* of this special collection. By this they mean the following. The outcome of their bargaining game is a coalition structure. At the equilibrium, the coalitions which are formed are chosen from the coalitions of the sequence  $S^*$  as follows. Given  $n$  choose the largest integer  $k^* \in S^*$ . This is  $k_{\hat{l}}^*$ . Eliminate this  $k_{\hat{l}}^*$  players. Now, choose the largest  $k^* \in S^*$  less than  $n - k_{\hat{l}}^*$  and so forth. Thus, their equilibrium consists of

specially selected elements of  $S^*$ . However, this coincidence of the results is only true for small values of  $n$  ( $n < 9$ ).

In contrast to the Cournot case, the results of the two games do not always coincide because with the contribution game, several coalitions can be formed. Because the payoff function is additive in the way the public good affects the utility (Ray and Vohra (2001) p.1374)  $k_1^*$  does not depend on  $n$  and forms in each sub-game in which there are  $k_1^*$  players left, whatever the coalitions which have been formed already. However, when there are more than  $k_1^*$  players left, the player designed by the protocol may find it more beneficial to form another coalition with some of the  $n - k_1^*$  players or to wait for the formation of another coalition, rather than to propose a larger coalition of size  $k_2^*$ . This explains, for example, the fact that when  $n = 12$  the grand coalition is an equilibrium in the game  $G$  and not in the game  $\Delta$ . Indeed, in the game  $\Delta$ , player 1, leaves the game independent, knowing that two coalitions of size 3 and 8 will be formed. In the game  $G$ , players do not have the possibility to form several coalitions. Therefore, player 1 compares her payoff when she is independent when the coalition of size 8 is formed and her payoff in the grand coalition. She proposes and forms the grand coalition.

## 6 Discussion

The introduction of this paper focuses on the theory of cartels since, like I do in this paper, this literature deals with situations in which symmetrical players have the possibility to reach a unique agreement, which generates positive externalities. However, it is part of a more general literature on coalition formation. In the framework of non-cooperative game theory, we can find a series of papers using extensive form bargaining games (Bloch (1996), Ray and Vohra (1999), Ray and Vohra (2001)) and other papers using normal form games (Hart and Kurz (1986), Yi

(1997)). In these papers, the outcomes of the game are coalition structures, the players are not always symmetric and the payoffs are not always decided by a fixed decision scheme. However, these games which are, from some points of view, rather general, have some underlying assumptions which are, in fact, quite restrictive. The aim of this paper has been to present a new framework which dispenses with these.

The literature on extensive form games of coalition formation made considerable progress on the generalisation of the game  $\Delta$  described in the paper. Bloch (1996), who first proposed this game, also proposed in the same contribution, a bargaining game in which players are not necessarily symmetric. He proves several results on the relationship between the equilibria of this extended game and different cooperative concepts. Ray and Vohra (1999) and (2001) drop the assumption of fixed payoff division and, in the symmetric case, they obtain as a result, the equal division which was assumed by Bloch.

However in all these games, the assumption is made that, once a coalition is formed, it is irreversible. This means that, as soon as a player has proposed a coalition which has been accepted by all the members, this coalition leaves the bargaining game. This assumption may have strong effects on the results. To understand this, consider the example of the game  $\Delta$  applied to Cournot payoff functions. Because of the positive externalities, which generate a free riding problem, players prefer to be independent when a large coalition is formed. Thus, the first players in the protocol leave the game as independents in order to induce the formation of the largest possible coalition of the remaining players. The assumption that players can leave the game and commit to remaining independent plays an important role here. Once the coalition has been

formed, no re-negotiation is possible, in spite of the fact that these independent players may have an incentive, ex post, to form another coalition<sup>2</sup>.

The advantage of normal form games is that there is no equivalent assumption. At the equilibrium, the different coalitions of a structure are not binding, but generated by self-enforcing choices. However in the game  $G$ , the assumption that only one coalition can be formed, in one sense, plays the same role as the assumption of irreversibility. In the same Cournot example, the independent players, by assumption, cannot form another coalition. The ideal answer would be to find a normal form game thereby dispensing with the irreversibility assumption but without the restriction to a single coalition. However, it is true that the problem will then arise of how to avoid the non-existence of equilibria.

Another assumption which is made in Bloch (1996) and Ray and Vohra (1999) and (2001) is the existence of a protocol which specifies the order in which players move. When players are symmetric, the protocol gives an explanation of asymmetric equilibria. However, when the players are different, we may want to distinguish the two sources of asymmetries: the intrinsic characteristics of the players and their position in a protocol. We saw that, in the game  $G$ , there is no specific protocol. Each strategy is a list of choices and each choice can be considered to be specific to each protocol of a simple extensive form game. Such a framework seems to be suitable for the analysis of agreements between heterogeneous players and a subject for future research is to find an extension of the game  $G$  to deal with such asymmetric cases.

One might be tempted to think that, if the objective is to avoid the assumptions used in extensive form games, it would be better to use more general normal form games such as those proposed

---

<sup>2</sup> For example, if a second coalition of size  $k_2$  is formed, each member's payoff in this new coalition will be  $\Pi^{k_2}(k_1, k_2) = \frac{1}{k_2(n - k_1 - k_2 + 3)^2}$  (see Bloch (1996), p. 113 for the general formula). When  $k_2 = n - k_1^*$ , it is easy



for example by Hart and Kurz (1987). Indeed, these authors propose two normal form games of formation of coalition structures. In both games, the strategy is, for each player, to give the coalition to which she wants to belong. In fact, these strategies are only wishes and the strategy profile a set of wishes, which are not necessarily consistent. Each game is then characterised by a rule mapping each strategy profile to a coalition structure. In the  $\Gamma$  rule, a coalition is formed if and only if all the members have chosen this coalition. In the  $\Delta$  rule, a coalition is formed if and only if all its members have chosen the same coalition, even if this coalition cannot be realised. Hence, the way the different wishes become consistent and generate a coalition structure is a rule of the game and has nothing to do with the rationality of the players.

The rules of the game  $G$  play also a role in the determination of which agreements are feasible and which ones will be realised. One definition says that an agreement is feasible if there are enough players who find it acceptable and the rule of efficiency says that only Pareto efficient feasible agreements can be achieved. However, these rules are not restrictive and it is difficult to imagine realistic situations in which they would not be satisfied. Moreover, the point is that they cannot oblige players to be in situations which they would rationally avoid. In particular, after a deviation, the new agreements which will be achieved depend not only on the new strategies of the deviating players, but also on the strategies of the non-deviating players. It is the Strong Nash equilibrium in restricted games which ensures that the complete strategy of each player is self enforcing.

Now, one might argue that it is too complicated to analyse conjectures about the reactions of non-deviating players in a normal form game. In that case, and if we do not want to revert to extensive form games, a possibility would be to use the framework of the theory of social situations developed by Greenberg (1990). The contributions to the theory of coalition formation by Chwe (1994) Mariotti (1997) or Xue (1998) are in this spirit. The players are able to foresee the

---

to check that,  $\Pi^I(k_1^*) < \Pi^{k_2}(k_1^*, k_2)$  as long as  $n < 41$ .

consequences of their deviations. The general idea is that, as in non-cooperative game theory, players are selfish and form non-binding agreements but they do not have to specify strategies for the whole game. This is in one sense a stability analysis. This framework, therefore, allows us to dispense with the definition of strategies, which is one of the main problems with extensive form games as Xue (1998) points out. However, there is an ongoing and unresolved discussion about the multiplicity of the consequences a player can forecast. The only way to solve this problem is then to impose behavioural assumptions.

In contrast to the latter approach, in this paper, I have dealt with the problem of conjectures by confining my attention to a normal form game and by focusing on the specification of the strategy spaces. It turns out that, in my restricted framework, these strategies are simple to define and have an intuitive interpretation.

## **7 Conclusion**

To conclude, the approach developed in this paper can be summarised and interpreted as follows. The objective is to find the solution to a bargaining problem involving the formation of an agreement, which generates positive externalities on the non-members. I do not use an extensive form bargaining game, which obliges one to give an explicit description of the process of agreement formation and involves making some restrictive underlying assumptions. Instead, I use a one shot normal form game in which each player's strategy is to say for each size of agreement whether it is acceptable or not. This means that the player will agree to sign any of the agreements that she designates as acceptable and remains alone otherwise. Next, an equilibrium refinement called the strong Nash equilibrium in restricted games is introduced, which guarantees that every one of the choices making up each strategy is self-enforcing. Under general payoff

functions, which exhibit positive externalities between the partners to the agreement and the independent players, I prove the uniqueness of the equilibrium, which I can also characterise.

This analysis has been conducted in a framework which is simplified by a number of assumptions such as, in particular, those specifying that only one agreement can be reached and that players are symmetric. The ideal answer would be to find a normal form game, thereby dispensing with the irreversibility and protocol assumptions but with asymmetric players and without the restriction to a single coalition. However, it is true that the problem will then arise of how to avoid the non-existence of equilibria.

## References

R. Aumann, 1959. Acceptable Points in General cooperative n-person games. In: Contributions to the Theory of Games IV, Princeton Univ. Press, Princeton, N.J.

d'Aspremont, C., J. Jaskold Gabszewicz, A. Jacquemin and J.A. Weymark, 1983. On the Stability of Collusive Price Leadership. Canadian Journal of Economics, pp. 17-25.

B.D Bernheim, B. Peleg and M.D. Whinston, 1987. Coalition-Proof Nash Equilibria, I Concepts. Journal of Economic Theory 42, 1-12.

B.D Bernheim, B. Peleg and M.D. Whinston, 1987. Coalition-Proof Nash Equilibria, II Applications. Journal of Economic Theory 42, 13-29.

F. Bloch, 1995. Endogenous Structures of Association in Oligopolies. Rand Journal of Economics 26, 537-556.

F. Bloch, 1996. Sequential formation of coalitions with fixed payoff division and externalities. Games and Economic Behavior, 14, 90-123.

M. S-Y Chwe, 1994. Farsighted Coalitional Stability. Journal of Economic Theory, 63, 299-325.

E. Diamantoudi, 20001. Stable Cartels Revisited. Working Paper, Department of Economics, University of Aarhus, 9.

J. Greenberg, 1990. The Theory of Social Situations. Cambridge: Cambridge University Press.

S. Hart and M. Kurz, 1983. Endogeneous Formation of Coalitions. Econometrica 51, 1047-1064.

M. Mariotti, 1997. A Model of Agreements in Strategic Form Games, Journal of Economic Theory, 74, 196-217.

D. Ray, R. Vohra, 1999. A Theory of Endogenous Coalition Structures. Games and Economic Behavior, 26, 286-336.

D. Ray, R. Vohra, 2001. Coalitional Power and Public Goods. *Journal of Political Economy*, 109, 1355-1384.

S. W. Salant, S. Switzer and R. J. Reynolds, 1983. Losses from Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium. *Quarterly Journal of Economics* 98, 185-199.

S. Thoron, 1998. Formation of a Coalition Proof Stable Cartel. *Canadian Journal of Economics* 31, 63-76.

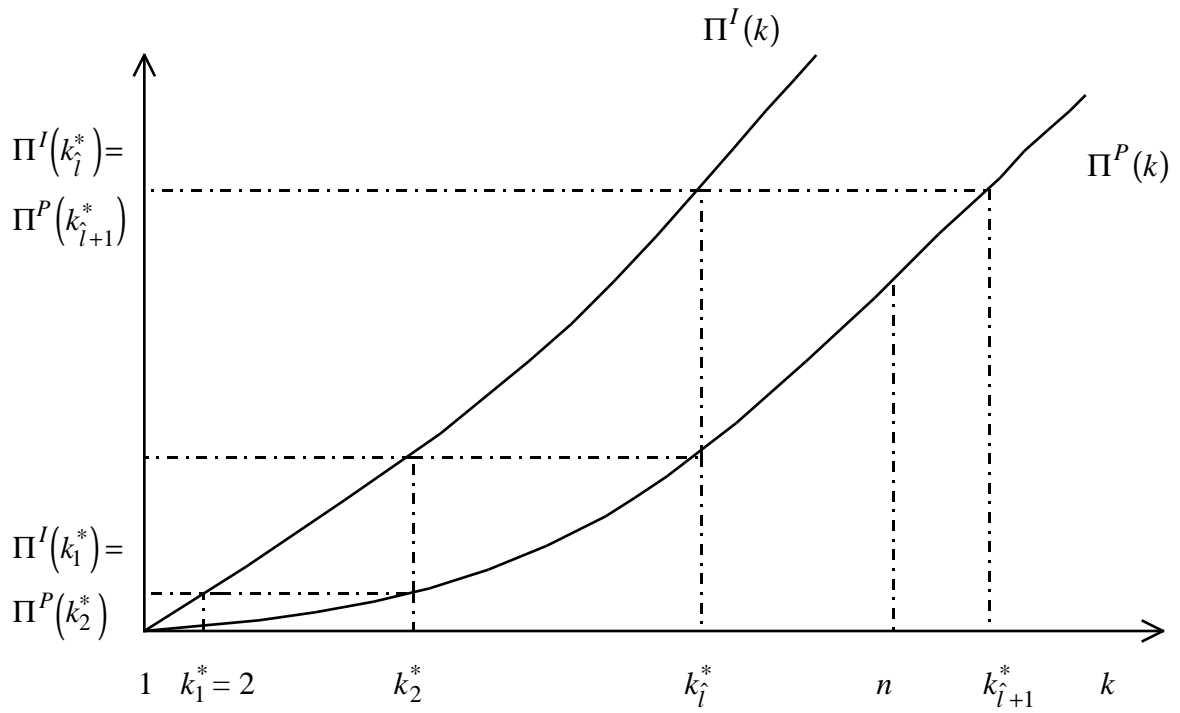
J. Von Neumann and O. Morgenstern, 1944. *Theory of Games and Economic Behaviour*. University Press, Princeton.

L. Xue, 1998. Coalitional Stability under Perfect Foresight. *Economic Theory*, 11, 603-627.

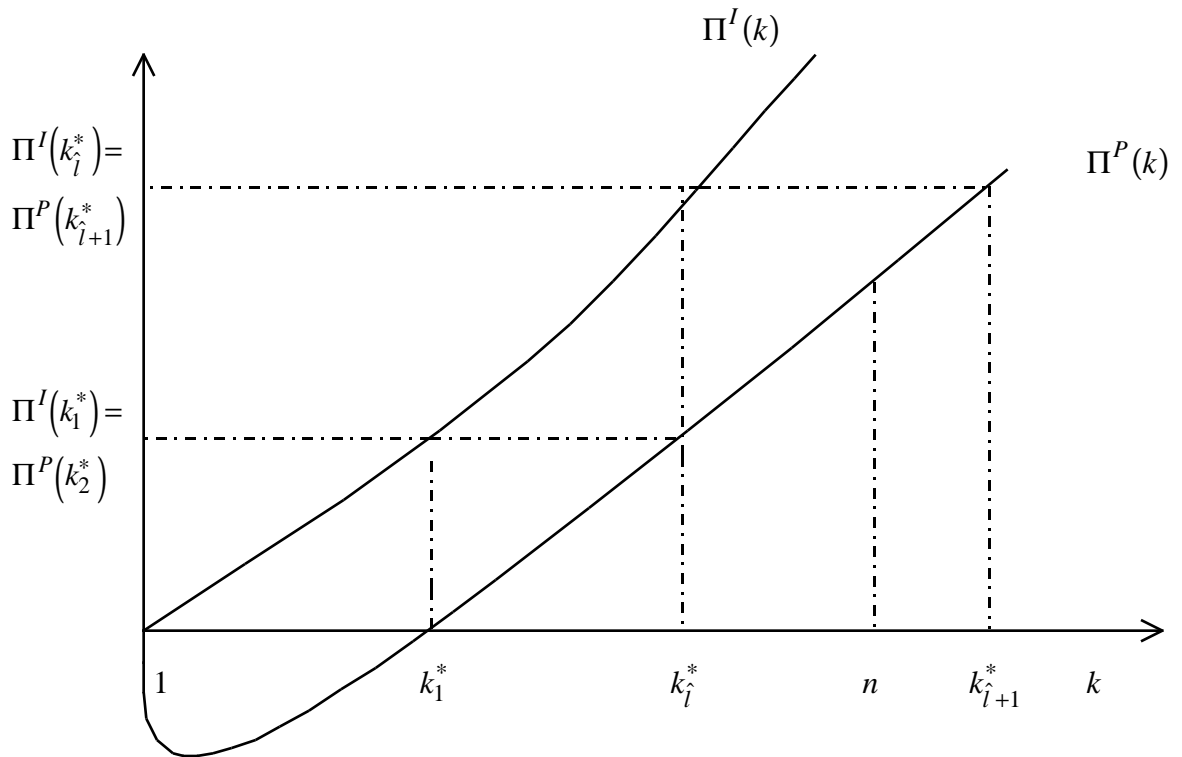
S.S. Yi, 1997. Stable Coalition Structures with Externalities. *Games and Economic Behaviour*, 20, 201-237.

**Table 1:** Strong Nash equilibrium in restricted games

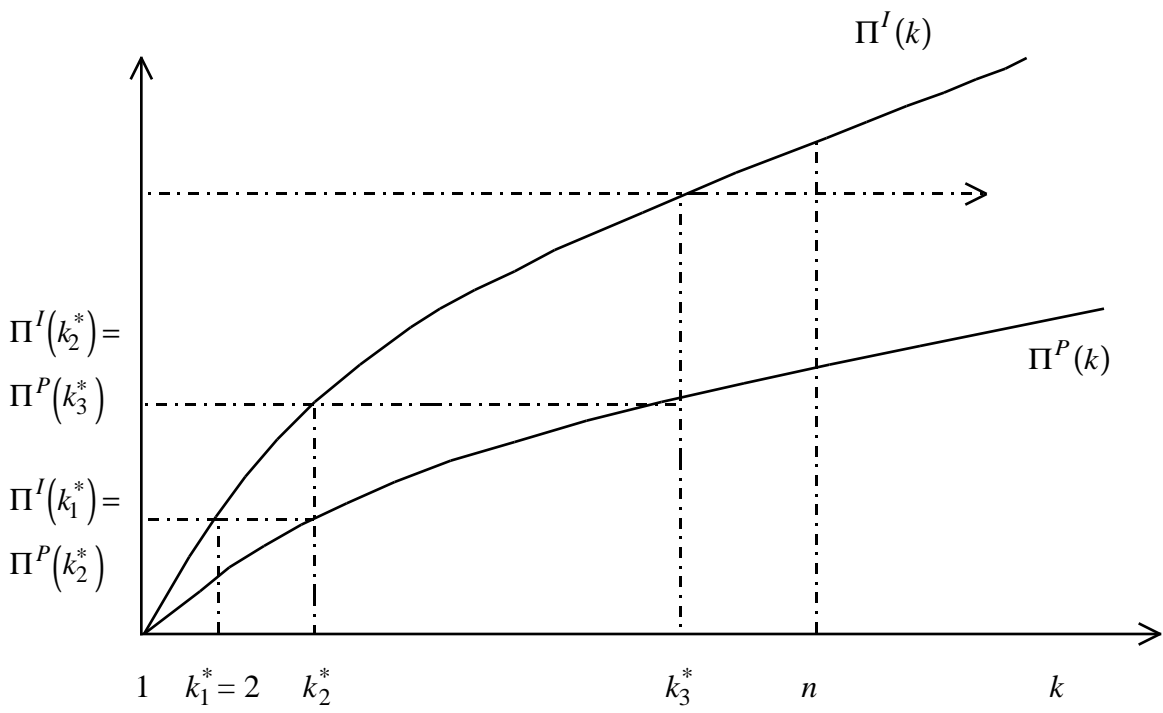
$k_1^*$	$k_2^*$	$k_3^*$	$k_4^*$	$k_5^*$	$k_6^*$	$k_7^*$	$k_8^*$	$k_9^*$
2	3	5	8	12	17	25	36	52
$n=2$	$n=3,4$	$5 \leq n < 8$	$8 \leq n < 12$	$12 \leq n < 17$	$17 \leq n < 25$	$25 \leq n < 36$	$36 \leq n < 52$	$52 \leq n < -$



**Figure 1**



**Figure 2**



**Figure 3**