Institutions Design for Managing Global Commons

by

Carlo Carraro

(University of Venice and FEEM)

Abstract

This paper provides some examples of how institution design affects the emergence of co-operative international agreements to manage global commons. The paper shows how different accession rules, minimum participation rules and negotiation rules affect a country's decision to sign or not to sign a treaty to protect a global common. The paper also analyses what would be the outcome of the negotiations when treaty design (e.g. the minimum participation rule or the negotiation agenda) is endogenised and strategically chosen by the negotiating countries.

Preliminary Version. April 2001

Prepared for presentation at the 4th FEEM-IDEI-INRA Conference on Property Rights, Institutions and Management of Environmental and Natural Resources, Toulouse, May 3-4, 2001.

The author is grateful to Carmen Marchiori, Francesca Moriconi, Sonia Oreffice and Jill Weinreich for their valuable assistance. The usual disclaimer applies. Corresponding address: Prof. Carlo Carraro, Department of Economics, University of Venice, San Giobbe 873, Venice, Italy. Tel: + 39 041 2574 166; Fax: + 39 041 2574 176; E-mail: ccarraro@unive.it

Institutions Design for Managing Global Commons

1. Introduction

In many cases, environmental quality is a public good. When the dimension of the environmental problem is global, as in the case of global commons (climate, ozone layer, biodiversity, etc.), there is no supra-national authority which can enforce the provision of the environmental good. In this case, sovereign countries must decide voluntarily whether or not to provide the public good and hence the level of emission abatement. In practice, countries negotiate on an international agreement which defines emission targets for each signatory and often also the way to achieve these targets.

Early contributions (Cf. Hardin and Baden, 1977) characterised the interaction among countries as a prisoners' dilemma, inevitably leading to the so-called "tragedy" of the global common property goods. However, in the real world, at the same time, a large number of international environmental agreements on the commons was signed, often involving sub groups of negotiating countries and sometimes involving transfers and other links with other policies (trade, technological cooperation, etc.).

How can we explain that some countries sometimes decide to sign an international environmental agreement even when they could enjoy the same environmental benefit by letting other countries to abate? In other words, when the environment is a public good, a country which does not abate achieves -- without paying any cost -- the same environmental benefit of signatories who decide to abate. Is this enough to conclude that there is no incentive for countries to sign an international agreement, i.e. that the outcome of the negotiation process will be a situation without any environmental cooperation?

The recent literature on international environmental agreements provides a negative answer to this last question. Indeed, many different papers have shown that cooperation can emerge even when the environment is a public good and when each country decides independently, voluntarily and without any forms of commitment and/or repeated interaction[1]. To get this result, of course more than two players must be involved in the environmental negotiation, and some conditions on the countries' reaction functions must be met.

To analyse the emergence of environmental cooperation, environmental economists have used a fairly new game-theoretic approach in which cooperation, i.e. the formation of coalitions, is the outcome of a non-

[1] The literature is now quite large. Let us just mention the works by Hoel (1991), Carraro and Siniscalco (1992, 1993), Barrett (1994), Heal (1994), Chander and Tulkens (1993, 1994). Surveys of this literature are provided by Tulkens (1996), Barrett (1997a) and Carraro (1997b).

cooperative strategic behaviour of the countries involved in the negotiation. This approach has been proposed both in games without spillovers (Le Breton and Weber, 1996, 1997) and in games with positive or negative spillovers (Bloch, 1996, 1997; Carraro and Siniscalco, 1993; Yi, 1997).[2]

The literature on environmental negotiations has emphasised the importance to model the decision process through which countries decide to join an environmental coalition as a non-cooperative game. The goal is to determine the so-called "self-enforcing agreements", i.e. agreements which are not based on the countries' commitment to cooperation. The game is assumed to be a two-stage game: in the first stage, countries decide non cooperatively whether or not to sign the agreement (join the coalition) given the burden-sharing rule which is adopted by the signatories countries; in the second stage, countries set their emission levels (their environmental policy) by maximising their welfare function given the decision taken in the first stage and the adopted burden-sharing rule.[3]

This theoretical literature has achieved some important conclusions (see Barrett, 1997b; Carraro and Siniscalco, 1998; Tulkens, 1998; Carraro, 1998a, for a few surveys):

- Even in the case in which there are positive spillovers (e.g. in the case of a public good provision) and even without any commitment to cooperation, countries may form a coalition, i.e. may decide to sign a treaty in order to cooperate to achieve a common target.
- This coalition is usually formed by a subgroup of the N negotiating countries (Barrett, 1994).
- The sometimes small initial coalition can be expanded by means of transfers or through "issue linkage", but only under some restrictive conditions, in particular when countries are symmetric (Carraro and Siniscalco, 1993, 1995; Botteon and Carraro, 1997a,b).
- the outcome of the two stage game described above crucially depends on the membership rules (open membership, exclusive membership, coalition unanimity, ...) that are adopted by the N negotiating countries (Carraro and Moriconi, 1998).

This latter result highlights the importance of properly designing the institutions which govern the management of global commons. Indeed, recent research stresses the strict dependence of the outcome of negotiations on global commons on the institutions and rules which are adopted by the negotiating countries. Some examples of these institutions and rules are the way in which new signatories are accepted (e.g. in the

---

2 It must be acknowledged that the first results on the non-cooperative formation of coalitions in the presence of positive spillovers can be found in the oligopoly literature on stable cartels. See D'Aspremont and Gabszewicz (1986), Donsimoni, Economides and Polemarchakis (1986).

3 Repeated interactions are usually ruled out for two reasons. First, because the threat mechanisms upon which the equilibrium is based are unrealistic in environmental situations. Second, because the grand coalition is one of the equilibria of the repeated game (given the usual conditions on the discount rate) and this is in contrast with a reality in which partial coalitions characterise the observed outcome of negotiations.

EU, new member countries are accepted only with the unanimous consensus of all existing members), the time sequence through which decisions are taken (e.g. should signatories decide simultaneously or can they adhere to the treaty sequentially?), a minimum participation level below which the treaty is not operational (in the Kyoto protocol at least 55 countries accounting for at least 55% of total emissions must sign the treaty), etc.

This paper has two goals. On the one hand, to show how different institution designs affect the structure of the equilibrium coalition (the number and identity of the signatory countries). On the other hand, to analyse how decisions about institutions are taken and whether there are the incentives to adopt the institutions that maximise the number of participants in the cooperative management of global commons.

Therefore, the next section summarises results concerning the dependence of the equilibrium coalition on the adopted accession rule. This section will show that a non trivial equilibrium coalition generally exists, but that its size varies largely as a function of the accession rule. Hence, the following section 3 attempts to endogenise countries' decision about the accession rule by adding a stage to the game. In the first stage, countries choose the accession rule, in the second stage they decide whether or not to sign the treaty and finally in the third stage they set their emission level. This section will show that, despite the incentive to free-ride also on the definition of the accession rule, at the equilibrium countries accept "to tight their hands" in the first stage in order to reduce the free-riding incentives in the second stage.

Section 4 endogenises another institutional aspect of the negotiation process on global commons. When defining the negotiation agenda, countries may decide to negotiate on one issue only, e.g. the cooperative management of a global common, or on several issues at the same time. In particular, the negotiation on an environmental problem can be linked to the negotiation on a different policy issue. When do countries have an incentive to link two different policy issues? Under what conditions? Section 4 will provide an answer to these questions by highlighting the trade-off between issue linkage and the benefits of separate negotiations.

Finally, section 5 addresses the issues of regional vs. global environmental treaties. It is indeed argued that the incentives to sign regional treaties, like in the case of trade blocs, are larger than the incentives to sign a single global agreement. Again, section 5 will explore this issue and provide conditions for regional agreements to emerge at the equilibrium.

A concluding section summarises the results of this paper.

---

[4] Other developments and advances can be found in Carraro (1997)

2. Accession Rules

Assume negotiations take place among N countries, $N \geq 3$, each indexed by $i=1, ..., n$. As said, countries play a two-stage game. In the first stage -- the *coalition game* -- they decide non-cooperatively whether or not to sign the agreement (i.e. to join the coalition). In the third stage, they play the non-cooperative Nash *emission game*, where the countries which signed the agreement play as a single player and divide the resulting payoff according to a given burden-sharing rule (any of the rules derived from cooperative game theory).[5]

This two-stage game can be represented as a *game in normal form* denoted by $\Gamma = (N, \{X_i\}_{i\in N}, \{u_i\}_{i\in N})$, where N is a finite set of players[6], X_i the strategy set of player i and u_i the payoff function of player i, assigning to each profile of strategies a real number, i.e. $u_i : \Pi_{i\in N} X_i \to R$. The payoff function is a twice continuously differentiable function.

A *coalition C* is any non-empty subset of N. A *coalition structure* $\pi = \{C_1, C_2, ... , C_m\}$ is a partition of the player set N, i.e. $C_i \cap C_j = \emptyset$ for $i \neq j$ and $\cup^m_{i=1} C_i = N$.

Since the formation of a coalition creates externalities, the appropriate framework to deal with this game is a *game in partition function form,* in which the payoff of each player depends on the entire coalition structure to which he belongs (Bloch, 1997; Ray and Vohra, 1996, 1997). This is why we convert the game in normal form into a game in partition function form.

Denote by Π the set of all feasible coalition structures. A *partition function* P: $\Pi \to R$ is a mapping which associates each coalition structure π with a vector in $\Re^{|\pi|}$, representing the worth of all coalitions in π. In particular, $P(C_i; \pi)$ assigns a worth to each coalition C_i in a coalition structure π. When the rule of payoff division among coalition members is fixed, the description of gains from cooperation is made by a *per-member partition function* p: $\Pi \to \Re^n$, a mapping which associates each coalition structure π with a vector of individual payoffs in $\Re^n$. In particular $p(C_i; \pi)$ represents the payoff of a player belonging to the coalition C_i in the coalition structure π[7].

Under suitable assumptions (A.2 and A.3 below), the second stage of the game can be reduced to a partition function (Yi, 1997; Bloch 1997). Therefore, the study of coalition formation consists of the study of the first

[5] This approach must be contrasted with the traditional cooperative game approach (e.g. Chander-Tulkens, 1993, 1995) and with a repeated game approach (Barrett, 1994, 1997b). Moreover, note that the regulatory approach often proposed in public economics is not appropriate given the lack of a supranational authority.

[6] We use the symbol N to denote both the number of players and the set of players. This is not a problem because we are going to assume that players are symmetric.

[7] Bloch (1997) denotes the per-member partition function by the term "valuation".

stage of the game, i.e. the negotiation process between the players. This negotiation process can be modelled either as a simultaneous game, in which all the players announce their strategic choice at the same time, or as a sequential one, in which each player can announce his strategy according to an exogenous rule of order. Let us first assume that:

A.1. All players decide simultaneously in all stages;[8]

Below we will briefly discuss how results change when decisions are taken sequentially. Moreover, it is helpful to assume that:

A.2. The second stage emission game has a unique Nash equilibrium for any coalition structure.

This assumption is necessary for the first stage of the game to be reduced to a partition function. However, in order to convert the strategic form into a partition function, the competition among the various coalitions has also to be specified. The common and perhaps the most natural assumption is that:

A.3. Inside each coalition, players act cooperatively in order to maximise the coalitional surplus, whereas coalitions (and singletons) compete with one another in a non-cooperative way.[9]

The partition function is then obtained as a Nash equilibrium payoff of the game played by coalitions and singletons. Formally, for a fixed coalition structure $\pi = \{C_1, C_2, \ldots, C_m\}$, let x^* be a vector of strategies such that:

$$\forall C_i \in \pi, \quad \sum_{j \in C_i} u_j\left(x^*_{C_i}, x^*_{N \setminus C_i}\right) \geq \sum_{j \in C_i} u_j\left(x_{C_i}, x^*_{N \setminus C_i}\right) \qquad \forall x_{C_i} \in \times_{j \in C_i} X_j$$

Then define:

$$P(C_i; \pi) = \sum_{j \in C_i} u_j(x^*).$$

Studying the issue of coalition formation by partition function games implies a limitation. The second-stage "reduction" procedure requires the grand-coalition to satisfy superadditivity, in the sense that the grand coalition should be able to achieve in terms of aggregate worth at least the sum of what is achievable under any coalition structure (Ray and Vohra, 1996, 1997; Bloch, 1997). Indeed, all strategies chosen by the

---

[8] By contrast, Barrett (1994) assumes that the group of signatories is Stackelberg leader with respect to non-signatories in the second stage emission game. In Bloch (1997) it is assumed that players play sequentially in the first stage coalition game.

[9] This assumption is equivalent to the γ-equilibrium of Chander and Tulkens (1995, 1997).

coalitions of any coalition structure can be replicated by the grand coalition. The superadditivity concerns just the grand coalition and it is not implied at the level of sub-coalitions, because of the presence of spillovers. This means that in any game of coalition formation, the grand coalition is always an efficient outcome whereas a fragmented coalition structure always results in an inefficient outcome.

In order to simplify the derivation of the partition function, we introduce a further assumption:

A.4. All players are ex-ante identical, which means that each player has the same strategy space in the third stage decision game.

This assumption allows us to adopt an equal-sharing payoff division rule inside any coalition, i.e. each player in a given coalition receives the same payoff as the other members[10]. Furthermore the symmetry assumption implies that a coalition C_i can be identified with its size c_i and a coalition structure can be denoted by $\pi=\{c_1,c_2, \ldots ,c_m\}$, where $\sum_i c_i = N$. As a consequence, the payoff received by the players depends only on coalition sizes and not on the identities of the coalition members. The per-member partition function (partition function hereon) can thus be denoted by $p(k; \pi)$, which represents the payoff of a player belonging to the size-k coalition in the coalition structure π. Finally, let us denote by $\pi = \{a_{(r)}, \ldots\}$ r size-a coalitions in the coalition structure π.

Recent developments of the theory of endogenous coalition formation (Cf. Bloch, 1997) have stressed the implications of allowing players to join different coalitions (see also Carraro, 1998b). However, in many cases, e.g. the case of global warming and climate change, the negotiating agenda focuses on a single agreement that players have to decide whether or not to sign. Here we assume that:

A.5. Players propose to sign a single agreement. Hence, those which do not sign cannot propose a different agreement. From a game-theoretic viewpoint, this implies that only one coalition can be formed, the remaining defecting players playing as singletons.

This assumption will be relaxed in section 5. There is a further a technical assumption which is useful to simplify the analysis of the coalition formation game:

A.6. Each player's payoff function increases monotonically with respect to the coalition size (the number of signatories in the symmetric case).[11]

---

[10] We consider the equal sharing rule as an assumption since it is not endogenously determined in the model. However Ray and Vohra (1996) provides a vindication for this assumption.

[11] This assumption excludes the possibility of "exclusive membership equilibria" where the group of cooperating players can refuse entry to a player which wants to join the coalition (Yi, 1997). Moreover, this assumption is useful to

This last assumption is quite natural in the case of negotiations on climate change control. Indeed, since climate is a public good, each country which finds it convenient to reduce its own GHG emissions provides a positive contribution to the welfare of all countries (both inside and outside the coalition). However, this condition can be relaxed to provide a more general analysis of the coalition formation game with different membership rule.

Using the above assumptions, it is possible to show that, as said in the Introduction, a non trivial equilibrium coalition $c^* \geq 2$ generally exists, even though the grand coalition $c = N$ is unlikely to form at the equilibrium (Barrett, 1994; Carraro and Siniscalco, 1993; Heal 1994; Hoel 1992). The existence and size of the equilibrium coalition depend on the intensity of the strategic interactions among countries, i.e. the slope of their reaction functions (Carraro and Siniscalco, 1993). Moreover, using the above assumptions, in sections 3 and 4 we will also be able to endogenise some of the features of negotiation process, namely the accession rules and the number of issues on which the negotiation focuses.

In this section, we would like highlight the strict dependence of the equilibrium coalition on some of the rules and institutions which are implicit in the above assumptions. For example, the negotiation process can be modelled either as a *simultaneous* game, in which all players announce at the same time their strategic choice -- as in A.1 --, or as a *sequential* one, in which each player can announce his strategy according to an exogenous order rule. Moreover, both a *linear* and a *circular* order of moves could be considered.

Accession rules are also very important to determine the size of the equilibrium coalition. Three different membership rules could be considered: the open membership rule, the exclusive membership rule and the coalition unanimity rule. In the *open membership game* (D'Aspremont *et al.* 1983; Carraro and Siniscalco, 1993; Yi and Shin, 1994)[12] each player is free to join and to leave the coalition without the consensus of the other coalition members. The negotiation process can be thought of as if each player announces a message. In this game all the players who announce the same message form a coalition. This membership rule thus implies that a coalition accepts any new player who wants to join it. By contrast, in the *exclusive membership game*[13] (Yi and Shin, 1994) or *game* Δ[14] (Hart and Kurz, 1983), each player can join a coalition only with the

---

distinguish the minimum coalition size endogenised in this paper from the minimum participation proposed in Black *et al.* (1992). In this latter paper, minimum participation is necessary to make the coalition profitable. In our paper, all non-trivial coalitions are profitable by assumption A.6. Hence, minimum participation may be chosen by players to increase the size of the stable coalitions (see below).

[12] The open membership game is "a game in which membership in a coalition is open to all players who are willing to abide by the rules of the coalition" (Yi and Shin, 1997).

[13] The exclusive membership game is a "a game in which the existing members in the coalition are allowed to deny membership to outsider players" (Yi and Shin, 1997)

consensus of the existing members, but he is free to leave the coalition. In this negotiation process, each player's message consists in a list of players with whom he wants to form a coalition. Then, players who announce the same list form a coalition, which is not necessarily formed by all players in the list. Finally, in the *coalition unanimity game*[15] (Chander and Tulkens, 1993; Yi and Shin, 1994; Bloch, 1997) or *game* Γ[16] (Hart and Kurz, 1983), no coalition can form without the unanimous consensus of its members. This implies that players are not free either to join the coalition or to leave it. In this negotiation process, players' messages consist in a list of players as in the previous one. However, if a coalition is formed it is necessarily composed of all players in the list and as soon as a player defects the coalition breaks up into singletons.

Another important element of our analysis is the type of conjecture formed by players when taking their decision about joining or leaving a coalition. We consider two cases. The usual *Nash conjecture* for which a player *i* takes his decision given the other players' decisions which are assumed not to change as a consequence of player *i*'s decision. This implies that when a player decides to leave a coalition, he conjectures that the other players will keep cooperating. By contrast, in the *rational conjecture* case, a player leaving a coalition can rationally predict the consequences of his decision and therefore how many players will then decide to join or leave the coalition (Ray and Vohra, 1996, 1997; Carraro and Moriconi, 1998).

Table 1 summarises the factors which need to be considered to determine the size of the equilibrium coalition. For each combinations of the three main characteristics – order of moves, membership rule, conjectures – we could also consider two additional factors:

(i) countries' reaction functions can be *orthogonal* or *negatively sloped.* The orthogonal case is particularly important in environmental games because it implies that free-riders enjoy the cleaner environment produced by the emission abatement carried out in some countries, but do not expand their emissions. By contrast, the non-orthogonal case is the one in which market mechanisms induce the so-called "leakage" problem, i.e. free-riders benefit twice from the other countries' cooperation, both because they get a cleaner environment and because they can increase their degree of exploitation of natural resources.

---

14 The game Δ is "a game in which the choice of a strategy by a player means the largest set of players he is willing to be associated with in the same coalition. Each set of all the players who chose the same C then forms a coalition (which may, in general , differ from C)" where C indicates a subset of players (Hart and Kurz, 1983).

15 The coalition unanimity game is a game in which "a coalition forms if and only if all potential members agree to join it" (Yi and Shin, 1994)

16 The game Γ is a game in which "each player chooses the coalition to which he wants to belong. A coalition forms if and only if all its members have in fact chosen it; the rest of the players become singletons" (Hart and Kurz, 1983).

(ii) the per-member partition function can be positively sloped and *monotonic* (with respect to the coalition size) -- as in A.6 --, or *humped-shaped*. In the monotonic case, and above a minimum coalition size c^m (see Carraro and Moriconi, 1998), the per-member partition function increases monotonically with respect to the number of signatories of the agreement. Formally, $p(c; \pi)$, where $\pi = \{c, 1_{(n-c)}\}$, is an increasing function of c for all $c>c^m$ (where $c^m=1$ with orthogonal free-riding). In the humped-shaped case, and above a minimum coalition size, there is an optimal size $c°<n$ at which the per-member partition function is maximised. Formally, $p(c; \pi)$, where $\pi = \{c, 1_{(n-c)}\}$, increases with c in the interval $(c^m,c°)$ and decreases with c for $c>c°$.

Table 1. Environmental Coalition Games. Taxonomy

	Open membership (monotonic or humped-shaped payoff)	Exclusive membership (humped-shaped payoff)	Coalition unanimity (monotonic or humped-shaped payoff)
Simultaneous game (orthogonal or negatively sloped reaction funct.)	Nash or rational conjectures	Nash or rational conjectures	Nash
Linear sequential game (orthogonal or negatively sloped reaction funct.)	Subgame perfect	Subgame perfect	-
Circular sequential game (orthogonal or negatively sloped reaction funct.)	Subgame perfect	Subgame perfect	-

Let us start from the simplest and most often analysed case in which decisions are simultaneous, membership is open and conjectures are the Nash's ones. The equilibrium is completely characterised by the following two properties, first derived in the cartel literature (D'Aspremont et al., 1983) and then often used also in the environmental literature (Carraro and Siniscalco, 1993; Barrett, 1994).

Profitability. A coalition is profitable if each co-operating player gets a payoff larger than the one he would obtain in the autarchic state, i.e. when no coalition forms. Formally:

(1) $$p(c;\pi) > p(1;\pi^S),$$

where $\pi=\{c,1_{(n-c)}\}$ and $\pi^S=\{1_n\}$, for all $i\in c$.

From eq. (1), c^m, the value of the minimal profitable coalition size, can be derived. This value depends on the strategic interaction between the coalition and the singleton players. In particular, with orthogonal free-riding any coalition size is profitable and c^m is simply two (Carraro and Siniscalco, 1992). By contrast, with non-orthogonal free-riding behaviour, the coalition has to reach a minimal size by which it can offset the damaging free-riders' action. This size is generally larger than two.

- Stability. A coalition is stable if it is both internally and externally stable. It is internally stable if no cooperating player is better off by defecting in order to form a singleton. Formally:

(2a) $$p(c;\pi) \geq p(1;\pi'),$$

where $\pi=\{c,1_{(n-c)}\}$ and $\pi'=\pi\backslash\{c\}\cup\{c-1, 1\}$for all $i\in c$.[19] It is externally stable if no singleton is better off by joining the coalition c. Formally:

(2b) $$p(1;\pi) > p(c';\pi'),$$

where $\pi=\{c,1_{(n-c)}\}$ and $\pi'=\pi\backslash\{c,1\}\cup\{c'\}$ and $c'=c+1$ for all $i\notin c$.

These two conditions define the Nash equilibrium of the first stage of the game. Let us denote by c* the size of the equilibrium coalition. The following proposition holds:

Proposition 1 (*Carraro and Siniscalco, 1993; Barrett, 1994*). The Nash equilibrium of a simultaneous single coalition game under the open membership rule is the following coalition structure:

- $\pi^* = \{c^*,1_{(n-c^*)}\}$, when $1\leq c^m\leq c^*<n$;
- the grand coalition c=n, when $c^*\geq n$;
- $\pi^S = \{1_n\}$, i.e. the singleton structure, otherwise;[20]

---

[19] We suppose that if a player is indifferent between joining the coalition or defecting, then he joins the coalition.

both when the payoff function is monotonic and when it is humped-shaped.
Proof: Carraro and Moriconi (1998).

The question is therefore the following. How does the equilibrium coalition change whenever the accession rule and/or the order of moves change? The variety of equilibrium coalitions which can be obtained by modifying the accession rules and/or the order of moves is discussed in Carraro, Marchiori and Moriconi (2001). They show that almost all coalition structures can be an equilibrium of the coalition game. In particular, coalition unanimity can sustain the grand coalition at the equilibrium (Chander and Tulkens, 1995, 1997), whereas exclusive membership (which is meaningful only when the payoff is humped-shaped) reduces the size of the equilibrium coalition, i.e. $c^{\circ} < c^*$. More generally, we can say that:

- The grand coalition $c=N$ is an equilibrium of the coalition game under coalition unanimity and under a circular order of moves with open membership.
- c^* is an equilibrium of the game under open membership both with simultaneous and with linear sequential choices. Sequential decisions enable us to identify the cooperating players (the first movers).
- Rational conjectures increase the size of the equilibrium coalition and make it possible the emergence of multiple equilibrium coalitions.
- Exclusive membership does not increase and is likely to reduce the size of the equilibrium coalition with respect to open membership ($c^{\circ} \leq c^*$).
- The presence of leakage is also likely to reduce the size of the equilibrium coalition and even to prevent any coalition to form at the equilibrium.
- The distinction between monotonic and humped-shaped payoff functions is relevant only under exclusive membership.

These conclusions show that the equilibrium coalition structure varies from the case of a small coalition (if it exists) to the case of the grand coalition. Further support to this conclusion derives from the introduction of transfers and/or issue linkage. These strategies can indeed be used to expand the number of signatories of the international agreement to manage global commons (Carraro and Siniscalco, 1998). As a consequence of this high variability of the equilibrium coalition structure, it is important to determine how countries agree on the rules and institutions that crucially affect the outcome of their negotiations. A first step into this direction will be accomplished in the next two sections.

---

[20] The symmetry assumption A.3 implies that there is not a single Nash equilibrium of the game. The number of Nash equilibria depends on the number of players, because, by symmetry, all subset of N formed by c^* players can be a NE of the game. We say that there is one Nash equilibrium because we refer to the typology of the equilibrium, rather than to the identity of players in the coalition.

3. Endogenous accession rules

As seen above, in the theoretical literature on the endogenous non-cooperative formation of coalitions the decision of countries to sign the treaty or not is usually modelled as a two stage game. In the first stage, countries non-cooperatively decide whether or not to sign, by anticipating the consequence of their decision on the economic variables under their control, whose value will be set in the second stage of the game.

Given the important effects of the membership rule on the equilibrium coalition shown in the previous section, we would like to add a new stage to the negotiation game. The game analysed in this section can be described as follows:

- in the first stage, countries choose the membership rule that will be used in the following stage;
- in the second stage, they decide whether or not to sign the treaty;
- in the third stage, they set the value of the economic variables under their control.

All decisions are to be taken non-cooperatively and simultaneously accordingly to assumption A.1. Moreover, assumption A.2-A.6 are also maintained in this section. In addition, we need an assumption about the way in which the membership rule is determined in the first stage of the game. Since this stage precedes the negotiation stage in which players decide whether or not to join the coalition, and since each player is independent, no player can be required to negotiate according to a membership rule that he does not agree upon. Hence, we assume that:

A.7. No player can be forced to accept a membership rule he does not agree upon. Hence, in case of disagreement on the membership rule, the less restrictive one, among those proposed by the N players, is adopted.

The main goal of this section is to endogenise the choice of the membership rule. However, a general treatment of this choice can hardly be provided. Hence, in the first stage of the game we will not consider a strategy space which includes all possible membership rules. We will rather focus on the so called α-rules, i.e. on rules which determine the minimum share α of the N negotiating countries which must sign the treaty for it to be effective.

This approach enables us to account for some of the most commonly assumed membership rules. Indeed, if $\alpha=0$, no restriction is introduced and a coalition can form whatever the number of its members. Hence, we are in the case of the so-called open membership rule. If instead $\alpha=1$, the coalition can form only if all countries decide to sign the treaty. Hence, we have the coalition unanimity rule. If $\alpha=0.55$, and countries are symmetric, we have the 55% rule introduced in the Kyoto protocol. Therefore, by endogenising α we can determine whether it is optimal for countries to choose the open membership rule, the coalition unanimity rule or any other rule defined by $\alpha \in [0,1]$.

Let us assume that c* is the size of the equilibrium coalition when $\alpha=0$ (open membership). This result was shown in the previous section and is the starting point of our analysis. It is obvious that any value of $\alpha' \leq \alpha^* = c^*/N$ does not modify the equilibrium coalition. If in the first stage players set $\alpha = \alpha' \leq c^*/N$, they ask for a coalition size for the treaty to be effective which is smaller than the size that would form anyway. Hence the condition is automatically satisfied and the equilibrium coalition remains c*.

If in the first stage players agree on a value of $\alpha^\circ > \alpha^* = c^*/N$, then in the second stage they have the choice either to form a coalition of size $c^\circ = \alpha^\circ N$, or not to form any coalition at all. Indeed, any coalition larger than c° is unstable, because any coalition larger than c* is unstable and $c^\circ > c^*$. Moreover, no coalition smaller than $c^\circ = \alpha^\circ N$ can be formed, by definition of α-rule. Given that, by assumption A.6:

(3) $$Q(c^\circ-1) > P(c^\circ) > P(\varnothing)$$

then countries prefer to form the coalition c°.

Hence, there is the following relationship between the membership rule decided in the first stage of the game and the equilibrium of the second stage coalition game:

Proposition 2 *(Carraro, Moriconi and Oreffice, 1999)*: The equilibrium of the coalition game for any value of $\alpha \in [0,1]$ is the coalition structure:

(4a) $$\pi^*(\alpha) = \{c^*, 1_{(N-c)}\} \text{ for any } \alpha\text{-rule such that } 0 \leq \alpha \leq \alpha^*$$

(4b) $$\pi^*(\alpha) = \{c=\alpha N, 1_{(N-c)}\} \text{ for any } \alpha\text{-rule such that } \alpha^* < \alpha \leq 1$$

This relationship between the value of α and the equilibrium of the coalition game, together with the incentive mechanisms which support the coalition c*, is crucial to determine the equilibrium value of α in the first stage of the game.

As said, the decision in the first stage of the game is taken independently and simultaneously by all players. As no player can be forced to accept a membership rule he does not agree upon, the less restrictive among the rules proposed by players is adopted. This rule must guarantee all players a payoff larger than the one they would get with an even less restrictive α-rule. This assumption is based on three ideas. First, if a player is ready to accept a restrictive α-rule, he can accept also a less restrictive one. Second, a commitment to an α-rule more restrictive than the one proposed by another player would not be credible because of the presence of positive spillovers. Third, in order to achieve a consensus on a given α-rule in the first stage, all players must prefer the equilibrium supported by that rule to the one that would emerge with a different less restrictive α-rule.

In the first stage, all players know that if no agreement is found on a value of α larger than $\alpha^* = c^*/N$, the equilibrium structure of the game would be $\{c^*,1_{(N-c)}\}$. A consensus on a value of α that leads to a coalition structure $\{c^\circ=\alpha^\circ N, 1_{(N-c)}\}$, $\alpha^\circ>\alpha^*$, can be found only if all players (weakly) prefer the coalition structure $\{c^\circ=\alpha^\circ N,1_{(N-c)}\}$ to the coalition structure $\{c^*,1_{(N-c)}\}$. Let $P(c) = p(c; \pi) - p(1; \pi^S)$. Notice that, in the presence of leakage, P(c) is positive for values of c above c^m. In a similar way, we can define the gain from free-riding on the coalition emission abatement. If a coalition c forms, a free-rider achieves $Q(c) = p(1; \pi) - p(1; \pi^S)$, where Q(c) is the free-riding function. Then, for $\{c^\circ=\alpha^\circ N,1_{(N-c)}\}$ to be preferred to $\{c^*,1_{(N-c)}\}$ by all countries, we must have:

(5a) $P(c^\circ) \geq P(c^*)$, i.e. all players who cooperate in $\{c^\circ,1_{(N-c)}\}$ and in $\{c^*,1_{(N-c)}\}$ must be better off in $\{c^\circ,1_{(N-c)}\}$.

(5b) $Q(c^\circ) \geq P(c^*)$, i.e. all players who are free-rider in $\{c^\circ,1_{(N-c)}\}$ but cooperators in $\{c^*,1_{(N-c)}\}$ must be better off when in $\{c^\circ,1_{(N-c)}\}$.

(5c) $P(c^\circ) \geq Q(c^*)$, i.e. all players who cooperate in $\{c^\circ,1_{(N-c)}\}$ but free-ride in $\{c^*,1_{(N-c)}\}$ must be better off in $\{c^\circ,1_{(N-c)}\}$.

(5d) $Q(c^\circ) \geq Q(c^*)$, i.e. all players who free-ride both in $\{c^\circ,1_{(N-c)}\}$ and in $\{c^*,1_{(N-c)}\}$ must be better off in $\{c^\circ,1_{(N-c)}\}$.

If these four conditions are satisfied, all players find it convenient to agree on the α-rule that sustains $\{c^\circ,1_{(N-c)}\}$. Notice that $c^\circ > c^*$ (otherwise the coalition structure is $\{c^*,1_{(N-c)}\}$). Hence, $P(c^\circ) > P(c^*)$ by Assumption A.6. Moreover, as previously explained, Q(c) is also monotonically increasing, which implies $Q(c^\circ) > Q(c^*)$ when $c^\circ > c^*$. Finally, the two results and the definition of c* imply:

(6) $$Q(c^\circ) > Q(c^\circ\text{-}1) > P(c^\circ) > P(c^*).$$

As a consequence, only (5c), i.e. $P(c^\circ) \geq Q(c^*)$, must be verified for all players to agree on the selection of an α-rule such that the equilibrium coalition structure becomes $\{c^\circ,1_{(N-c)}\}$.

Is this condition feasible? The answer is certainly positive. Given the monotonicity of the functions P(c) and Q(c), it is indeed always possible, if N is sufficiently large, to find a coalition c° large enough to satisfy $P(c^\circ)$

$\geq Q(c^*)$. Suppose then that c° such that $P(c°) \geq Q(c^*)$ exists. The monotonicity of P(c) also implies that (5c) is satisfied for all $c'' \geq c°$. However, recall that members of any coalition c" larger than c* have an incentive to defect from c" because $Q(c''-1)>P(c'')$ and that this incentive is anticipated by players when setting α in the first stage of the game. Let $\alpha°$ be the solution of $P(c°=\alpha N) = Q(c^*)$. If a player in the first stage proposes a value α' larger than $\alpha°$, then one of the players that would belong to the coalition $c' = \alpha' N$ has an incentive to propose a value $\alpha'-1/N$, because he prefers to be a free-rider with respect to the coalition c'-1 which is sustained by $\alpha'-1/N$ (we are using (4b) here).

Hence, the following proposition holds.

Proposition 3 *(Carraro, Moriconi and Oreffice, 1999).* Given (4b), which defines the equilibrium of the coalition game for any value of $\alpha \in [c^*/N,1]$, and (2a)(2b) which defines the critical value $c^* \geq 1$, the α-rule $\alpha°$, where $\alpha°$ is the solution of $P(c°=\alpha N) = Q(c^*)$, is an equilibrium. As a consequence the equilibrium coalition structure is $\{c°=\alpha° N, 1_{(N-c)}\}$

In other words, in the first stage players agree to choose the smallest value of α which make all of them better off with respect to the coalition structure $\{c^*, 1_{(N-c)}\}$. This is a Nash equilibrium because no player has an incentive to propose a value of α larger than $\alpha°$ -- a different player would immediately propose a smaller one – and no player has an incentive to propose a value smaller than $\alpha°$ -- in this case at least one player would disagree (see Figure 1).

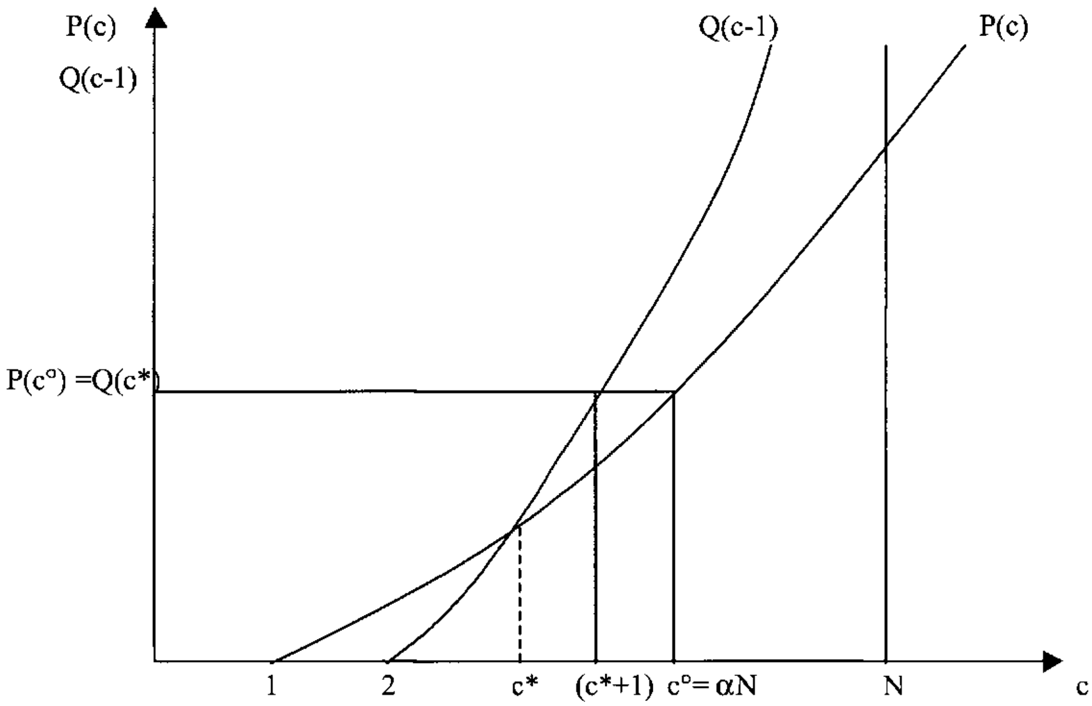
By setting $\alpha > c^*/N$, players reduce their freedom to free-ride in the second stage and increase their equilibrium payoff. However, this mechanism does not necessarily lead to the grand coalition, because there is an incentive "to undercut" the value of α in order to exploit the larger benefits arising from free-riding on the cooperators' efforts. The balance of these two mechanisms lead to the equilibrium value of α defined in Proposition 2.

The above results can be used to explain why non-cooperative players may decide to join a coalition and why they prefer to introduce an α-rule that partially reduces the negative effects of free-riding on the size of the equilibrium coalition.

---

[21] A more detailed analysis is provided in Carraro, Marchiori, Oreffice and Moriconi (2001).

Figure 1. Equilibrium coalition with endogenous accession rule.



4. Endogenous issue linkage

Another idea often proposed to reduce the incentives to free-ride on global commons is issue linkage. The basic idea behind issue linkage is to design an agreement in which countries do not negotiate only on one issue (e.g. the environmental issue), but are forced to negotiate on two issues, e.g. the environmental one and another interrelated (economic) issue. For example, Barrett (1995, 1997) proposes to link environmental negotiations to negotiations on trade liberalisation, Carraro and Siniscalco (1995, 1997), Katsoulacos (1997) propose to link environmental negotiations to negotiations on R&D co-operation, Mohr (1995) and Mohr et al. (1998) propose to link climate negotiations to international debt.

The idea of issue linkage was originally formulated by Folmer *et al.* (1993) and Cesar and De Zeeuw (1996) to solve the problem of asymmetries among countries. The intuition is that some countries gain on a given issue, whereas other countries gain on a second one. By linking the two issues, the agreement in which the countries decide to co-operate on both issues may become profitable to all of them.

However, the idea of issue linkage can also be used to offset the free riding incentives that characterise many environmental (or other public good) agreements. Suppose there is no profitability concern (either because countries are symmetric or because a transfer scheme is implemented to make the agreement profitable to all countries). Then, the equilibrium number of signatories of the linked agreement may be larger than the equilibrium number of signatories of the agreement whose equilibrium number of signatories is small because of the presence of strong incentives to free-ride (Carraro and Siniscalco, 1995). Of course this depends on the relative economic weights of the two agreements and on their nature. In particular, to be successful, an agreement concerning a (quasi) public good should be linked to an agreement concerning a (quasi) club good.

Let us consider an example. Suppose the environmental negotiation is linked to the negotiation on R&D co-operation[22], which involves an excludable positive externality (club good) and increases the joint coalition welfare. In this way, the incentive to free-ride on the benefit of a cleaner environment (which is a public good fully appropriable by all countries) is offset by the incentive to appropriate the benefit stemming from the positive R&D externality (which is a club good fully appropriable only by the signatories). The latter incentive can stabilise the joint agreement, thus also increasing its profitability because countries can reap both the R&D co-operation and the environmental benefit (this second benefit would be lost without the linkage).[23]

This idea is exploited in Carraro and Siniscalco (1995, 1997), Katsoulacos (1997), Botteon and Carraro (1998). Carraro and Siniscalco (1995, 1997) show that issue linkage may be a powerful tool to increase the number of signatories of an environmental agreement. For example, if developed countries on the one hand increase their financial and technological co-operation with developing countries, and on the other hand make this co-operation conditional on the achievement of given environmental targets, then a number of countries is likely to be induced to join the environmental coalition, i.e. to sign a treaty in which they commit themselves to adequate reductions in their emission growth. Katsoulacos (1997), which accounts for information asymmetries, provides additional support to the conclusion that issue linkage can be very effective in guaranteeing the stability of an environmental agreement. By contrast, Botteon and Carraro (1998), where asymmetric players are also considered, show that issue linkage may be counter productive because of potential conflicts among asymmetric players on the "optimal" members of the R&D club.

All the above-mentioned papers assume that issue linkage has been chosen by the negotiating countries.[24] Then they analyse the effectiveness of issue linkage in increasing the equilibrium number of signatories of

---

[22] See Carraro-Siniscalco (1997) for a full presentation of the model.

[23] A detailed proof of this result is in Carraro-Siniscalco (1997).

[24] An interesting exception is the paper by Tol *et al.* (1999).

the joint agreement (with respect to the environmental agreement). However, the decision of linking two economic issues should also be considered as a strategic choice that players undertake. A game therefore describes the incentives to link two issues. And this game could also be characterised by free-riding. Hence, even if issue linkage increases the number of signatories -- and therefore the amount of public good provided (e.g. emission abatement) -- it may not be an equilibrium outcome. The crucial question is therefore the following: do players have an incentive to link the negotiations on two different issues instead of negotiating on the two issues separately? Namely, is the choice of issue linkage an equilibrium of the game in which players decide non-cooperatively whether or not to link the negotiations on two different economic issues?

To answer these questions, let us consider a game in which, in the first stage, players decide whether to link two issues on which they are trying to reach an agreement or to negotiate on two separate agreements. If they decide not to link the two issues, in the second stage they can decide whether or not to sign either one or both separate agreements. If they decide in favour of issue linkage, in the second stage they decide whether or not to sign the linked agreement. Finally, in the third stage they set the value of their policy variables.

Two cases could be considered. One in which the benefits accruing to the signatories of one of the two separate agreements are perfectly or almost perfectly excludable (co-operators provide a club good) and one in which the degree of excludability is low. Moreover, the choice made in the first stage of the game could be analysed under two voting rules. One is unanimity, because the choice of issue linkage can be considered as a negotiation rule whose determination precedes the beginning of actual negotiations and which therefore should be taken with the consensus of all countries involved in the negotiations. A second one is majority voting, because it may be relevant to analyse how a less restrictive voting rule increases the chances of adopting issue linkage. In this section, we focus on the case of imperfect excludability and unanimity voting (consistently with the previous section). A general analysis is provided in Carraro and Marchiori (2001).

Let c_u^* identify the size of the equilibrium coalition when issue linkage is adopted; c_u^* is an equilibrium iff:

(7a) $$P_a(c_u^*) + P_t(c_u^*) \geq P_a(0) + P_t(0)$$

(7b) $$P_a(c_u^*) + P_t(c_u^*) \geq Q_a(c_u^*-1) + Q_t(c_u^*-1)$$

(7c) $$P_a(c_u^*+1) + P_t(c_u^*+1) < Q_a(c_u^*) + Q_t(c_u^*)$$

From (7a) it is clear that, if the individual agreements are profitable, then the linked agreement is also profitable. But not vice-versa. This is why, as explained above, issue linkage has been proposed to solve the profitability problem (Cesar and De Zeeuw, 1996).

Let us define the structure and the payoffs of the linkage game. If players decide to link the two issues and negotiate on a joint agreement, the equilibrium payoffs are:

(8a) $P_u(c_u^*) = P_a(c_u^*) + P_t(c_u^*)$ for a signatory of linked agreement;

(8b) $Q_u(c_u^*) = Q_a(c_u^*) + Q_t(c_u^*)$ for a free-rider.

If instead players prefer not to link the two issues, they decide whether or not to participate in two different agreements. In this case they obtain at the equilibrium:

(9a) $P_a(c_a^*) + P_t(c_t^*)$ if they decide to cooperate on both issues;

(9b) $P_a(c_a^*) + Q_t(c_t^*)$ if they cooperate in the "a-agreement" (the public good agreement), but they free ride on the "t-agreement" (the club good agreement);

(9c) $Q_a(c_a^*) + P_t(c_t^*)$ if they cooperate in the "t-agreement", but free ride on the other issue;

(9d) $Q_a(c_a^*) + Q_t(c_t^*)$ if they free-ride on both issues.

Hence, without linkage, there are four "types" of countries, where the identity of countries is irrelevant because of symmetry. The structure of the linkage game and its payoffs are summarised in Figure 2.

Under what conditions will players choose to link the two issues and negotiate on a joint agreement? In order to answer the above question, we need to identify the size of the linked coalition with respect to the coalition size of the two separate negotiations. Indeed, if by linking the two negotiations all benefits of the larger one are not lost and players achieve in addition the benefits from expanding the smaller coalition, then it is obvious that linkage is a profitable option for all players. However:

Proposition 4 *(Carraro and Marchiori, 2001)*: At the equilibrium, $c_u^* \leq c_t^*$, i.e. the number of players who participate in the joint agreement is always smaller than or equal to the number of players who participate in the "t agreement" linked to the public good agreement.

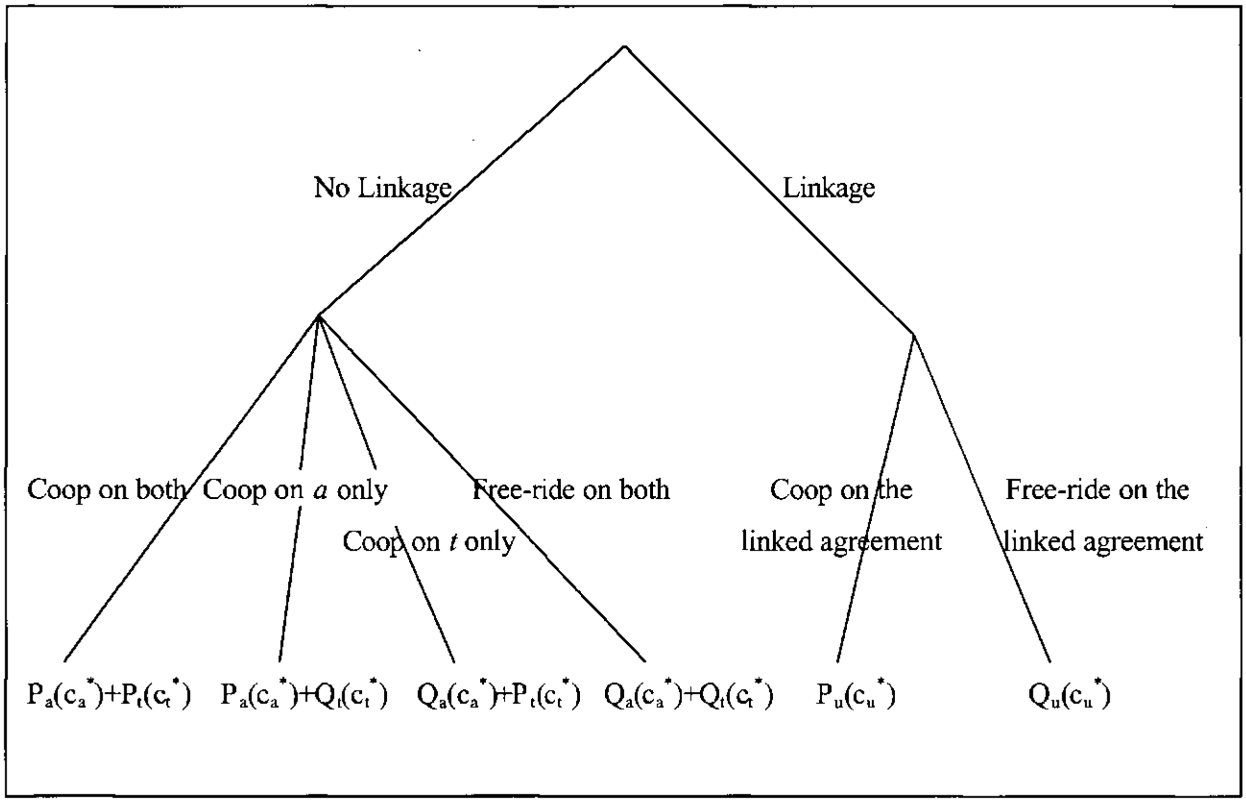
Proof: The joint agreement is internally stable if $P_u(c_u^*) \geq Q_u(c_u^*-1)$, i.e. if :

(10) $Q_a(c_u^*-1) - P_a(c_u^*) \leq P_t(c_u^*) - Q_t(c_u^*-1)$

For $c_u^* > c_a^*$, the left hand side of (10) is positive because there is an incentive to free-ride on the "a-agreement" for all $c > c_a^*$. This implies that the right had side is also positive, i.e. $P_t(c_u^*) > Q_t(c_u^*-1)$. Therefore, when countries co-operate on the "t-agreement", there is still an incentive to enter the coalition. Hence, c_u^* must be smaller than or equal to the equilibrium coalition size, i.e. $c_u^* \leq c_t^*$.

Q.E.D.

Figure 2. The structure of the game with endogenous issue linkage



Hence, assuming $c_a^* < c_u^*$, because otherwise issue linkage would be meaningless, Proposition 4 implies:

(11) $$c_a^* < c_u^* \leq c_t^* \quad \text{and} \quad c^\circ_t \leq c^\circ_u .$$

Given this ranking, the conditions for issue linkage to be an equilibrium of the three stage game can easily be derived by comparing the payoff shown in Figure 2. The result of this analysis is summarised by the following proposition:

Proposition 5: Assume c_i^*, i=a,t,u, $2\leq c_i^*<n$ are the equilibrium coalitions of the a (e.g. public good), t (e.g. club good) and linked agreement respectively. If the cooperative payoffs $P_i(c_i)$, i=a,t,u, increase monotonically with the coalition size, then issue linkage is players' equilibrium choice under unanimity voting iff:

$$P_a(c_u^*) - Q_a(c_a^*) > Q_t(c_t^*) - P_t(c_u^*) \tag{12}$$

Proof: See Carraro and Marchiori (2001).

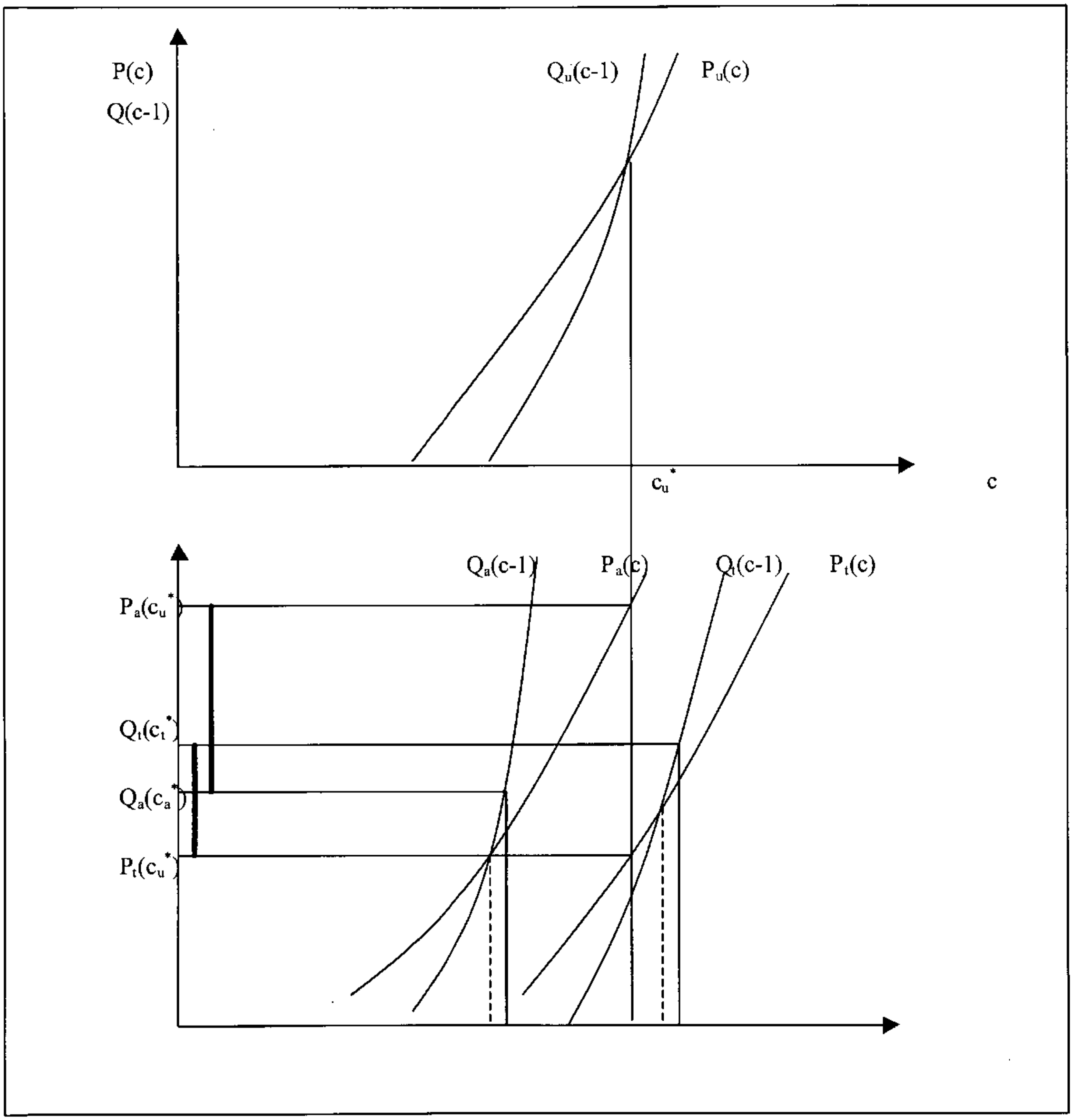
How can condition (12) be interpreted? $Q(c_t^*) - P_t(c_u^*)$ is the loss from reducing the coalition on the "t-agreement" from c_t^* to c_u^* (Proposition 4 shows that $c_t^* \geq c_u^*$). This loss can be written as $Q_t(c_t^*) - P_t(c_u^*) = [Q_t(c_t^*) - Q_t(c_u^*)] - [P_t(c_u^*) - Q(c_u^*)]$ where the first term represents a free-rider's loss when they get less benefits from a smaller coalition, whereas the second term represents the excess benefit of co-operation when $c_u^* < c_t^*$. Notice that $Q(c_t^*) - P_t(c_u^*) > 0$, because it is also equal to $[Q_t(c_t^*) - P_t(c_t^*)] + [P_t(c_t^*) - P_t(c_u^*)]$, where $[Q_t(c_t^*) - P_t(c_t^*)] > 0$ and $[P_t(c_t^*) - P_t(c_u^*)] > 0$, if $c_t^* > c_u^*$ and if $P_t(c_t)$ is monotonic.

$P_a(c_u^*) - Q_a(c_a^*)$ is the benefit (e.g. the environmental benefit) which a free-rider on the "a-agreement" achieves when joining the expanded coalition. It can be also written as $[P_a(c_u^*) - P_a(c_a^*)] - [Q_a(c_a^*) - P_a(c_a^*)]$ where the first term is the increased gain which a cooperator on the "a-agreement" achieves from expanding the coalition, whereas the second term is a free-rider's relative gain when a coalition c_a^* forms. The positivity of $Q_t(c_t^*) - P_t(c_u^*)$ implies that (12) holds if $P_a(c_u^*) - Q_a(c_a^*)$ is also positive, i.e. if *the increased gain which a cooperator on the "a-agreement" achieves from expanding the coalition is larger than a free-rider's relative gain when a coalition c_a^* forms.*

This is only a necessary condition. The sufficient condition says that *the increased gain which a cooperator on the "a-agreement" (e.g. a signatory of an environmental agreement) achieves from expanding the coalition from c_a^* to c_u^*, plus the excess benefit of co-operation on the "t-agreement" when $c_u^* < c_t^*$, must be larger than a free-rider's relative gain when a coalition c_a^* forms plus the loss that a free-rider suffers because of the smaller spillovers from the reduced co-operation on the "t-agreement".*

Proposition 5 can be generalised to the cases in which the function $P_t(c_t)$ may not be monotonic as in Carraro and Siniscalco (1997) or in which the negotiation on the t-agreement leads to the grand coalition as in Yi (1997) (in this latter case $Q(c_t^*=n)$ may be smaller than $P_t(c_t^*=n)$). These generalisations are proposed in Carraro and Marchiori (2001).

Figure 3. Geometric representation of the condition for issue linkage to be an equilibrium of the game



4. Regional vs. global treaties

Another idea sometimes proposed to reduce the free-riding incentives which undermine the possibility of achieving a global agreement on global commons is to give negotiating countries the freedom to sign different agreements at a regional level, in a way which is similar to the one adopted for trade agreements. Indeed, it is argued that a regional cooperative management of global commons could more easily be achieved than a global management of global commons.

Of course this raises a trade-off between the greater efficiency of a global agreement, which is however unlikely to be signed, and the possibly greater incentives to sign several regional agreements.

Let us explore this trade-off. Hence, let us consider the case in which assumption A.5 which imposes a single coalition is relaxed, i.e. countries are allowed to form as many coalitions as they prefer. This is equivalent to assume that any coalitions can be free-ridden by non singleton coalitions.

If multiple coalitions can form, then the outcome of the game is a coalition structure $\pi = \{c_1, ..., c_m\}$. However, the determination of the equilibrium coalition structure is not as easy as in the previous sections, for several reasons. First, in a coalition game with externalities, the worth of any coalition depends on the behaviour of the complement set. In a single coalition game, free-riders can behave solely as singletons and thereby the worth of the lone coalition is easily computable and the stability of a coalition structure coincides with the stability of the lone coalition. By contrast, in a multiple coalition game the complement set behaviour can be defined in several different ways. As a consequence, a coalition has not a unique worth and any notion of stability cannot be simply referred to a coalition, but must take into account the whole coalition structure to which such a coalition belongs. Second, the feasible coalition structures increase significantly when the number of players increases.

For these reasons we will not be able to analyse all cases (combinations of different order of moves, conjectures and membership rules) that were considered in section 2, but we will restrict ourselves to one illustrative case where decisions are taken simultaneously with Nash conjectures and open membership.[25] Moreover, the equilibrium structure will be identified by a stability condition (see below) which corresponds, as in section 2, to the Nash equilibrium of the coalition game. Hence, the stability condition is a strictly non-cooperative concept based on the assumption of Nash conjectures. We does not consider other solution concepts, e.g. those proposed by Bloch (1996), Ray and Vohra (1996, 1997), Chew (1994), Mariotti (1997), which make use either of some forms of commitment or of some forms of non-Nash conjectures.

---

25 We are not going to discuss the exclusive membership rule because it is relevant only when the payoff function is humped shaped, a case which does not add much to the understanding of the formation of multiple coalitions. We are

Let us start by extending the definitions provided for the single coalition case to the case in which multiple coalitions can form.

Positive spillovers. In any coalition structure, if coalitions merge to form a larger coalition, other coalitions not affected by the change are better off. Formally: consider two coalition structures π and π' and a coalition c_i such that $c_i \in \pi$, $c_i \in \pi'$, $c_j, c_k \in \pi'$ and $c_j \cup c_k \in \pi$. Then, $p(c_i; \pi) > p(c_i; \pi')$ for all players belonging to c_i.

Hence, the partition function is such that the payoff of a player belonging to a given coalition c_i is larger the larger the size and the lower the number of coalitions formed by players not belonging to c_i. This has an important implication. When players who do not join c_i decides to behave as singletons, they give rise to the worst possible complement structure for the coalition c_i, i.e. the minimax one. The single coalition game thus constitutes a benchmark for the multiple coalition one. For any fixed coalition size, the single coalition structure defines the minimum worth such a coalition can obtain.

Free riding. In any coalition structure, small coalitions have higher per-member payoffs than big coalitions. Formally: for any coalition structure π and any two coalitions c_i and c_k in π, $p(c_i; \pi) < p(c_k; \pi)$ if and only if $c_i > c_k$.

As in the single coalition game, two different free-riding behaviour patterns (orthogonal and non-orthogonal) can be singled out. Let us assume monotonicity (an increase of cooperators increases the payoff of all players).

Profitability. A coalition structure π is profitable if any coalition c_i in π is profitable. A coalition $c_i \in \pi$ is profitable if any cooperating player belonging to c_i gets a payoff larger than the one he would get in the singleton structure. Formally:

(13) $$p(c_i; \pi) \geq p(1; \pi^S)$$

for all players in the coalition c_i and all coalitions in π.

Let us recall that the payoff of a player belonging to c_i increases with the size of the other coalitions in π, which provide a positive externality. Geometrically, this implies that the payoff function in the multiple

---

not going to discuss the coalition unanimity rule because it leads trivially to the grand coalition if the profitability condition is satisfied. For the same reason, we do not discuss the case of circular sequential moves.

coalition case is shifted upward by the presence of more than one non-trivial coalition. Hence, the following conclusions are straightforward.

- If a coalition is profitable in the single coalition game, it is profitable in the multiple coalition game too.
- If c^m denotes the minimal coalition size above which a coalition becomes profitable in a single coalition game, then this minimal size becomes equal or smaller in a multiple coalition game with more than one non trivial coalition.

Let us also characterise the equilibrium of the game by extending the notion of Nash stability previously provided to the case of multiple coalitions. The internal and external stability conditions of section 2 are again necessary conditions, because it is still possible for any cooperating player to deviate to form a singleton, and for a singleton to join a coalition. However, these two conditions are no longer sufficient to characterise the equilibrium coalition structure. Indeed, with multiple coalitions, players choose both whether or not to join a coalition (as in the single coalition game) and which coalition to join. This is why a further stability condition on the entire coalition structure has to be added. We will call this latter condition "intracoalition stability".

Nash Stability (*Yi and Shin, 1994*). A multiple coalition structure π is stable if each coalition $c_i \in \pi$ is internally stable, externally stable and intracoalition stable. It is *internally stable* if no cooperating player would be better off by leaving the coalition to form a singleton. Formally:

$$p(c_i; \pi) > p(1; \pi') \tag{14a}$$

for all players in the coalition c_i and all coalitions in π, where $\pi' = \pi \setminus \{c_i\} \cup \{c_i-1, 1\}$. It is *externally stable* if no singleton would be better off by joining any coalition belonging to the coalition structure π. Formally:

$$p(1; \pi) > p(c_i; \pi') \tag{14b}$$

for all players who do not belong to c_i or to any other non-trivial coalition in π, where $\pi' = \pi \setminus \{c_i, 1\} \cup \{c_i+1\}$. It is *intracoalition stable* if no player belonging to c_i would be better off by leaving c_i to join any other coalition $c_j \in \pi$. Formally:

$$p(c_i; \pi) > p(c_j+1; \pi') \tag{14c}$$

for all players in the coalition c_i and all coalitions in π, where $\pi' = \pi \setminus \{c_i, c_j\} \cup \{c_i-1, c_j+1\}$. As in the single coalition case, these three conditions define the NE of the game.

As said above, the presence of multiple coalitions shifts upward the profitability function of a country joining a given coalition (with respect to the single coalition case). A similar conclusion cannot be proposed for the stability function. Indeed, when the concentration of the external coalition structure increases both the worth of belonging to a coalition and the value of behaving as a singleton increase. Suppose there is one coalition. Then a second coalition forms. Players in the first coalition receive a benefit (more abatement), but a benefit is also received by those players who do not join any coalition, and keep enjoying a cleaner environment without cooperating on emission abatement. There is a case in which the comparison is easy. If there is no leakage, i.e. country's best-reply functions are orthogonal, then cooperators in the first coalition and free-riders receive the same benefit from the second coalition abatement. Moreover, this benefit is additive, i.e. it is possible to write the payoff of each player in the game as the payoff resulting from his choice of joining (or not-joining) a coalition c_i, plus the benefit (positive externality) produced by the other coalitions in π. In the sequel, we will focus mostly on this case, which helps clarifying the mechanism of multiple coalition formation. However, we will briefly discuss also the case of non-orthogonal free-riding.

Let us normalise to zero the non-cooperative payoff $p(1, \pi^S)$. In the case of orthogonal free-riding, it is then possible to write a player's payoff function as $p(c_i; \pi) = P(c_i) + \Sigma_{j\neq i}Q(c_j)$, where the index j denotes all coalitions belonging to the coalition structures π other than c_i. In words, a player in c_i achieves his cooperative payoff plus the payoff from free-riding on the other coalitions that may form. In the case in which only two coalitions c_i and c_k form, the payoff would be $p(c_i; \pi) = P(c_i) + Q(c_k)$. From the definition of free-riding, $p(c_i; \pi) \leq p(c_k; \pi)$ if and only if $c_k \leq c_i$. Hence, $P(c_i) + Q(c_k) \leq P(c_k) + Q(c_i)$, i.e.

$$(15) \qquad P(c_i) - P(c_k) \leq Q(c_i) - Q(c_k)$$

This is equivalent to say that a player who moves to a larger coalition increases his payoff (by monotonicity) but the incremental benefit is lower than the incremental gain he would achieve by free-riding on the larger coalition.

Moreover, the case of orthogonal free-riding enables us to determine exactly the behaviour of the stability function for a given coalition c_i belonging to π. Indeed, if a player leaves a coalition c_i to be a singleton he compares $p(c_i; \pi)$ and $p(1; \pi')$, $\pi' = \pi \setminus c_i \cup (1, c_i\text{-}1)$, where:

$$(16) \qquad p(c_i; \pi) - p(1; \pi') = [P(c_i) + \Sigma_{j\neq i}Q(c_j)] - [Q(c_i\text{-}1) + \Sigma_{j\neq i}Q(c_j)] = L(c_i)$$

which coincides with the stability function previously defined. If $L(c_i)$ is non-negative, no player would like to leave the coalition to be a singleton. The same function $L(c_i)$, with the reversed sign, describes the incentive for a singleton to join the coalition c_i. Hence, if $L(c_i)$ is non-positive, no singleton would like to join the coalition c_i.

If a player leaves a coalition c_i to join a different coalition c_k, he compares $p(c_i; \pi)$ and $p(c_k+1; \pi'')$, $\pi'' = \pi \backslash c_i, c_k \cup (c_i-1, c_k+1)$, where:

$$(17) \qquad p(c_i;\pi) - p(c_k+1;\pi'') = [P(c_i) + \Sigma_{j\neq i}Q(c_j)] - [P(c_k+1) + Q(c_i-1) + \Sigma_{j\neq i,j\neq k}Q(c_j)]$$
$$= P(c_i) - P(c_k+1) + Q(c_k) - Q(c_i-1)$$
$$= L(c_i) - L(c_k+1)$$

First, notice that $L(c_i) \geq 0$ is no longer a sufficient condition for the absence of free-riding. Indeed, $p(c_i; \pi) - p(c_k+1; \pi'')$ can be negative even when $L(c_i)$ is positive. Second, by definition of free-riding, a rational player decides to leave c_i to join c_k only if $c_k < c_i$, which implies $c_k+1 \leq c_i$. Hence, $p(c_i; \pi) - p(c_k+1; \pi'') = L(c_i) - L(c_k+1) \leq 0$ because the stability function L(.) is decreasing when free-riding is orthogonal Hence, intracoalition stability is achieved only when $L(c_i) = L(c_k+1)$ which implies $c_i = c_k+1$. Therefore, at the equilibrium, coalition sizes can differ by at most one player.

We can therefore conclude that a NE coalition structure must satisfy the following restrictions:

- it contains only coalitions whose size is smaller than or equal to c^* (internal stability);
- it cannot contain simultaneously a singleton and a coalition smaller than c^* (external stability);
- coalition sizes differ by at most one player (intracoalition stability), i.e. at most two types of coalitions emerge at the equilibrium.

These remarks lead to the following proposition.

Proposition 6 *(Carraro and Moriconi, 1998).* The stable coalition structures of a simultaneous multiple coalition game with Nash conjectures and open membership are:

- $\pi^S = \{1_n\}$, i.e. the singleton structure, when $c^m > c^*$.
- $\pi^x = \{c^x{}_{(q)}, (c^x - 1)_{(p)}\}$ when $c^m \leq c^*$,

  with $c^x \leq c^*$, where p and q are two integers such that $c^x q + (c^x - 1)p = n$.

Proof: See Carraro and Moriconi (1998).

For example, with four players, two size-2 coalitions form (unless $c^*=4$ in which case there is only the grand coalition). If $c^*=3$, a size-3 coalition could form. However, $P(c^*)$ -- the payoff of a player in a size-3 coalition – is not larger than $P(c^*-1)+Q(c^*-1)$ -- the payoff of a player in a size-2 coalition when a second size-2 coalition forms. Indeed, $P(c^*)-Q(c^*-1) \leq P(c^*-1)$, i.e. $L(c^*) \leq P(c^*-1)$, because $L(c^*) \approx 0$ by definition of stable coalition and $P(c^*-1) > 0$ by profitability. If $c^*=2$, the conclusion that two size-2 coalitions form follows immediately.

With five players, if $c^*=5$, then the grand coalition forms. If $c^*=4$, a size-3 and a size-2 coalition form. Indeed, using the same argument as above, $P(c^*) \leq P(c^*-2)+Q(c^*-1)$, where the left-hand side denotes the payoff a player achieves by leaving c^* and forming a size-2 coalition. This is larger than $Q(c^*-1)$ -- the payoff a player achieves by leaving c^* and behaving as a singleton. If $c^*=3$, again a size-3 and a size-2 coalition form. Indeed, $P(c^*)+Q(c^*-1)$ -- the payoff a player achieves when he belongs to a size-3 coalition and another size-2 coalition forms – is not lower than $Q(c^*-1)+Q(c^*-1)$ – the payoff he would achieve by free-riding on two size two coalitions (of course, it does not make sense to leave a size-3 coalition to join a size-2 coalition that then is formed by three players).

With six players, if $c^*=6$, the grand coalition forms. If $c^*=5$, {4,2} is preferred to {5,1} because $P(c^*) \leq P(c^*-3)+Q(c^*-1)$, i.e. it is more profitable to form a small coalition and free-ride on the larger one. However, the equilibrium is not {4,2}. Indeed, by (10), a player in the size-4 coalition prefers to leave and join the size-2 coalition. We have that $P(c^*-1)+Q(c^*-3)$ – the payoff a player achieves in the size-4 coalition – cannot be larger than $P(c^*-2)+Q(c^*-2)$ – the payoff he achieves by moving to the size-2 coalition. Indeed, $P(c^*-1)+Q(c^*-3) \leq P(c^*-2)+Q(c^*-2)$, i.e. $P(c^*-1) - P(c^*-2) \leq Q(c^*-2) - Q(c^*-3)$. Moreover, there is no incentive to free-ride on two size-3 and size-2 coalitions, i.e. the structure {3,2,1} is not an equilibrium, because $P(c^*-2)+Q(c^*-2) \geq Q(c^*-3)+Q(c^*-2)$ by definition of c^*. Hence, the structure {3,3} is an equilibrium of the coalition game. However, this is not the only equilibrium, because the structure {2,2,2} can be shown to be an equilibrium as well. If $c^*=4$, again {4,2} is not an equilibrium because $P(c^*) \leq P(c^*-2)+Q(c^*-1) \leq P(c^*-1)+Q(c^*-1)$. Hence, {3,3} is again an equilibrium of the coalition game. It can be shown that {2,2,2} is also an equilibrium. Finally, if $2 \leq c^* \leq 3$, the application of Proposition 7 is straightforward. Two equilibrium structures exist, i.e. either three size-2 coalitions or two size-3 coalitions can emerge at the equilibrium.

When the number of players is larger than 6, a similar analysis confirms the conclusions of Proposition 6. For example, if $n=7$ and $c^*=4$, then the NE coalition structures can be {4,3} and {3,2,2}. If $n=8$, and $c^*=7$, the NE coalition structures are {4,4}, {3,3,2}, {2,2,2,2}. This multiplicity of equilibria can be reduced by moving to a sequential order of moves or by using the Coalition Proof Nash equilibrium concept (see Carraro and Moriconi, 1998).

The intuition behind Proposition 6 could be phrased as follows. A player prefers to free-ride cooperatively rather than as a singleton if the coalition he joins is not larger than c^*. Indeed, in this case he gets the double benefit of free-riding on the other coalitions abatement and of cooperating with a small number of players. By definition of c^*, this second benefit is larger than the gain from free-riding as a singleton.

The main implication of Proposition 6 is that more than one non-trivial coalition generally characterises the equilibrium coalition structure (provided that $c^*>1$, $n \geq c^*$ and $n>3$). However, this does not necessarily imply that emission abatement is higher when multiple coalitions can form. Consider again the case in which n= 8 and c*=7. In the single coalition case the unique Nash stable coalition has size equal to c*=7. Only one player decides to free-ride. By contrast, in the multiple coalition case, the Nash stable coalition structures are {4,4}, {3,3,2}, {2,2,2,2}. It is quite likely that the abatement carried out by a coalition formed by 7 players (and no leakage) abates more than four small size-2 coalitions and probably even more than 2 size-four coalitions.

This last conclusion is a consequence of the crucial result – an implication of intracoaliton stability – that at the equilibrium coalition sizes cannot differ by more than one member. Hence, even when c* is large and close to n, the multiple coalition equilibrium structure tends to be quite dispersed, i.e. several small coalitions form at the equilibrium. For example, if n= 20 and c* = 18, coalition structures such as {18,2}, {17,3}, {16,4} ... {10,5,5}, etc. cannot be equilibria because they do not satisfy intracoalition stability. The most concentrated Nash equilibrium structure is {10,10}, but {5,5,5,5}, {4,4,4,4,4}, {3,3,3,3,3,3,2}, {2,2,2,2,2,2,2,2,2,2}, etc. can also form.

A last important question has to be addressed. How do the above results extend to the case of non-orthogonal free-riding? In particular, in the presence of leakage, can we still conclude that the equilibrium coalition structure is characterised by more than one non-trivial coalition? We are unable to provide a general answer to this question. However, we would like to describe the two forces whose interaction characterises the equilibrium of the game. As said above, when free-riding is orthogonal, there is an incentive to form more than one coalition because cooperators in the first coalition do not modify their choice when a second coalition forms. Hence, members of the second coalition get the benefit from free-riding on the first coalition plus the benefit of cooperation in their own coalition (which must be smaller than or equal to c*). When free-riding is non-orthogonal, the abatement strategies of all players change as a consequence of the formation of the second coalition. Hence, for example, the first coalition may reduce its size or even collapse. In this case, members of the second coalition (still smaller than c*) gets the benefit of cooperation but may loose part or all of the benefit from free-riding on the first coalition. If this latter loss is larger than the benefit of cooperation, a second coalition is not going to form.

6. Conclusions

The previous sections analysed the effects of different institution designs on the possibility of achieving an effective cooperative management of global commons. Using a non-cooperative game-theoretic approach, we have shown that institutions matters and that different rules modify the incentives to sign a cooperative agreement to control polluting emission or the use of natural resources. In particular, there exist rules and institutions which can support a large participation in an international agreement to manage global commons.

The next question which we addressed was therefore whether there are the incentives to adopt these rules, namely under what conditions countries are likely to adopt rules which offset their free-riding incentives. This paper has analysed, again within a game-theoretic framework, some of these possible rules and the related incentives to adopt them. In particular, as far as accession rules are concerned, we showed that all countries find it profitable to reduce their freedom by introducing a minimum participation level which reduces free-riding thus inducing countries to form a larger equilibrium coalition.

We also analysed the incentives to link negotiations on global commons to other economic negotiations. Our results highlight under what conditions it is profitable for all countries to link two different negotiations in a way which increases the signatories of the agreement on global commons.

Finally, we discussed the incentives to sign an environmental agreement when countries are free to choose whether to cooperate in a single global agreement or in several regional agreements. The answer is that incentives lead countries to form several agreements (coalitions) even though the resulting equilibrium may not be the optimal one from an environmental viewpoint.

References

Barrett, S. (1994), "Self-Enforcing International Environmental Agreements", *Oxford Economic Papers*, 46, 878-894.

Barrett, S. (1995), "Trade Restrictions in International Environmental Agreements", mimeo, LBS.

Barrett, S. (1997a), "Heterogeneous International Environmental Agreements", in C. Carraro, ed., International Environmental Agreements: Strategic Policy Issues, E. Elgar, Cheltenham.

Barrett, S. (1997b), "Towards a Theory of International Cooperation" in C. Carraro and D. Siniscalco, eds., New Directions in the Economic Theory of the Environment, Cambridge University Press, Cambridge.

Bloch, F. (1997), "Noncooperative Models of Coalition Formation in Games with Spillovers", in C. Carraro and D. Siniscalco, eds., New Directions in the Economic Theory of the Environment, Cambridge University Press, Cambridge.

Botteon, M. and C. Carraro (1997a), "Burden-Sharing and Coalition Stability in Environmental Negotiations with Asymmetric Countries", in C. Carraro, ed., International Environmental Agreements: Strategic Policy Issues, E. Elgar, Cheltenham.

Botteon, M. and C. Carraro (1997b), " Strategies for Environmental Negotiations: Issue Linkage with Heterogeneous Countries", in H. Folmer and N. Hanley, eds., Game Theory and the Global Environment, E. Elgar, Cheltenham.

Carraro, C. (1997a), ed., International Environmental Agreements: Strategic Policy Issues, E. Elgar, Cheltenham.

Carraro, C. (1997b), "Environmental Conflict, Bargaining and Cooperation" forthcoming in J. van den Bergh, Handbook of Natural Resources and the Environment, E. Elgar, Cheltenham.

Carraro, C. and C. Marchiori (2001), "Endogenous Issue Linkage in International Negotiations", prepared for the Vth meeting of the Coalition Theory Network, Louvain-la-Neuve, 25-26.1.2001

Carraro, C., Marchiori, C. and F. Moriconi (2001), "Stable Coalitions", mimeo, University of Venice.

Carraro, C., Marchiori, C., Moriconi, F. and S. Oreffice (2001), "Endogenous Minimum Participation in International Environmental Treaties", mimeo, University of Venice.

Carraro, C. and F. Moriconi (1998), "Endogenous Formation of Environmental Coalitions", presented at the Coalition Theory Network Workshop on "Coalition Formation: Applications to Economic Issues", Venice, 8-10.1.98.

Carraro, C., Moriconi, F. and S. Oreffice (1999), "α-Rules and Equilibrium Endogenous Coalitions", presented at the 4th CTN Workshop on "Coalition Formation", Aix en Provence, 8-10 January 1999 and at the EAERE Annual Conference, Oslo, 24-26 June 1999.

Carraro, C. and D. Siniscalco (1992), "The International Protection of the Environment: Voluntary Agreements among Sovereign Countries", in Dasgupta, P and K.G. Maler, The Economics of Transnational Commons, Clarendon Press, Oxford.

Carraro, C. and D. Siniscalco (1993), "Strategies for the International Protection of the Environment", *Journal of Public Economics*, 52, 309-328.

Carraro, C. and D. Siniscalco (1995), "Policy Coordination for Sustainability: Commitments, Transfers, and Linked Negotiations", in I. Goldin and A. Winters, eds., The Economics of Sustainable Development, Cambridge University Press, Cambridge.

Carraro, C. and D. Siniscalco (1997), "R&D Cooperation and the Stability of International Environmental Agreements", in C. Carraro, ed., International Environmental Agreements: Strategic Policy Issues, E. Elgar, Cheltenham.

Carraro, C. and D. Siniscalco (1998), "International Environmental Agreements. Incentives and Political Economy", in *European Economic Review*, 42, 561-572.

Cesar, H. and A. De Zeeuw (1996), "Issue Linkage in Global Environmental Problems", in A. Xepapadeas, ed., Economic Policy for the Environment and Natural Resources, E. Elgar, Cheltenham.

Chander, P. and H. Tulkens (1995), "A Core-Theoretical Solution for the Design of Cooperative Agreements on Trans-Frontier Pollution", *International Tax and Public Finance*, 2, 279-294.

Chander, P. and H. Tulkens (1997), "The Core of an Economy with Multilateral Environmental Externalities", *International Journal of Game Theory*, 26, 379-401.

Chew, M.S. (1994), "Farsighted Coalitional Stability", *Journal of Economic Theory*, 54, 299-325.

D'Aspremont, C.A., A. Jacquemin, J.J. Gabszewicz and J. Weymark (1983) "On the Stability of Collusive Price Leadership", *Canadian Journal of Economics*, 16, 17-25.

Folmer, H., van Mouche, P. and S. Ragland (1993), "Interconnected Games and International Environmental Problems", *Environmental Resource Economics*, 3, 313-335.

Heal, G. (1994), "The Formation of Environmental Coalitions", in Carraro, C., ed., Trade, Innovation, Environment, Kluwer Academic Publisher, Dordrecht.

Hoel, M. (1992), "International Environmental Conventions: the Case of Uniform Reductions of Emissions", *Environmental and Resource Economics*, 2, 141-159.

Hoel, M. (1994), "Efficient Climate Policy in the Presence of Free-Riders", *Journal of Environmental Economics and Management*, 27, 259-274.

Katsoulacos, Y. (1997), R&D Spillovers, R&D Cooperation, Innovation and International Environmental Agreements, in C. Carraro, ed., International Environmental Agreements: Strategic Policy Issues, E. Elgar, Cheltenham.

Mariotti, M. (1997), "A Model of Agreements in Strategic Form Games", *Journal of Economic Theory*, 73, 128-139.

Nordhaus, W. (1999), "Global Public Goods", IDEI Annual Lecture, Toulouse, 14 June 1999.

OECD (1993), The Cost of Cutting Carbon Emissions: Results from Global Models, Paris.

Ray, D. and R. Vohra (1996), "A Theory of Endogenous Coalition Structure", mimeo, Brown University.

Ray, D. and R. Vohra (1997), "Equilibrium Binding Agreements", *Journal of Economic Theory*, 73, 30-78.

Tulkens, H. (1998), "Cooperation versus Free-riding in International Environmental Affairs: Two Approaches", in H. Folmer and N. Hanley, eds., Game Theory and the Global Environment, E. Elgar.

Yi, S. (1997), "Stable Coalition Structures with Externalities", *Games and Economic Behaviour*, 20, 201-23.