'CORE'

REPRINT

1339

COOPERATION VERSUS FREE-RIDING IN INTERNATIONAL ENVIRONMENTAL AFFAIRS: TWO APPROACHES

Henry TULKENS

CENTER FOR OPERATIONS RESEARCH & ECONOMETRICS
UNIVERSITE CATHOLIQUE DE LOUVAIN

VOIE DU ROMAN PAYS, 34
B -1348 LOUVAIN-LA-NEUVE
B E L G I U M

2. Cooperation versus free-riding in international environmental affairs: two approaches[1]

Henry Tulkens

1. INTRODUCTION

This chapter is about a controversy regarding the feasibility, and as a consequence the likelihood, of cooperation among countries on issues of transfrontier pollution. I want to contrast two theses, a pessimistic one and an optimistic one. Both of them are based on concepts rooted in economic analysis, and both of them claim additional support from game theory. Nevertheless they reach opposing conclusions. It is thus a challenging task to try to disentangle the arguments used on each side, in order to see whether the two theses can be reconciled or are intrinsically antagonistic.

The structure of the chapter is as follows. Before entering into the dispute, I think it is important to remind the reader, in Section 2, of how economic analysis shows that cooperation raises a severe problem in international environmental affairs, and what the logical structure of that problem is. Section 3 then contains a summary presentation of the two theses. Section 4 (which is the heart of the chapter) is devoted to a systematic comparison of their respective characteristics, and to a search for a conceptual framework for reconciling the two approaches. Section 5 concludes with considerations on the two basically different notions of 'stability' for coalitions that are at stake.

2. THE UNDERLYING ECONOMIC–ECOLOGICAL MODEL AND THE QUESTIONS RAISED

I briefly[2] remind the reader of the structure of the economic model, which is common to the two theses. A set N of countries, indexed by $i = 1, 2, \ldots, n$, share a common environmental resource. For each country, the function $u_i(x_i, z)$ describes national preferences over the consumption of some private good

($x_i \geq 0$) and of some environmental good ($z \leq 0$).[3] The function is assumed to be of the quasilinear form $u_i(\cdot) = x_i + v_i(z)$, with v_i concave and increasing. Define $\pi_i = (\partial u_i/\partial z / \partial u_i/\partial x_i) \geq 0$ as country i's marginal willingness to pay (in commodity x) for the environmental good. Furthermore, let $y_i = g_i(p_i)$ be country i's production function, linking[4] its output $y_i \geq 0$ of the private good with its emissions $p_i \geq 0$ of pollutant in the environment, and assume $\gamma_i = dy_i/dp_i > 0$ up to some maximum value p_i^o, and zero above it. The derivative γ_i is then naturally interpreted, when taken to the left, as the country's marginal cost (in y_i) of abating its emissions.

The 'transfer function' $z = -\Sigma p_i$ specifies how the pollutant emissions of all countries are diffused and transformed by ecological processes into the ambient quantity z. And finally, the private good is assumed to be transferable[5] between the countries, in amounts denoted as T_i (< 0 if given away by country i, > 0 if received by it).

For the *economic–ecological* system so described, we have:

Definition 1: A *feasible state* is a vector

$$(x, y, p, z, T) \equiv (x_1, \ldots, x_n; y_1, \ldots, y_n; p_1, \ldots, p_n; z; T_1, \ldots, T_n)$$

such that:

$$\begin{aligned} \forall i, \quad x_i &= y_i + T_i \\ y_i &= g_i(p_i) \\ \textstyle\sum x_i &= \textstyle\sum y_i \\ z &= -\textstyle\sum p_i. \end{aligned}$$

Notice that the first three constraints imply $\Sigma T_i = 0$.

Definition 2: A *non-cooperative equilibrium* in the sense of Nash is a feasible state $(\bar{x}, \bar{y}, \bar{p}, \bar{z}, \bar{T})$ such that:

$$\begin{aligned} &\forall i, (\bar{x}_i, \bar{p}_i) \text{ maximizes } x_i + v_i(z) \\ &\quad \text{s.t.} \quad y_i = g_i(p_i) \\ &\quad\quad p_i + z = -\sum_{j \neq i} \bar{p}_j, \end{aligned}$$

letting $\bar{T}_i = 0 \; \forall_i$.

Definition 3: An *internationally efficient state* (or, for short, an *international optimum*) is a feasible state $(x^*, y^*, p^*, z^*, T^*)$ that maximizes

$$\sum_{i \in N} [x_i + v_i(z)].$$

The well-known fact – readily established from first order conditions – that the non-cooperative equilibrium is not an international optimum suggests that environmental efficiency at the world level can only be achieved through some form of cooperation among the countries involved. This is the source of the economists' motivation for interpreting and/or designing international treaties as instruments towards world efficiency.

But what should be the contents for such a treaty? and which countries the signatories? If all countries are convinced of the need to cooperate, an 'efficient' treaty would naturally specify the joint abatement policy corresponding to the internationally optimal emissions vector $(p_1^*, \ldots, p_n^*)$ derived above.[6] For some countries, however, the emission policy p_i^* may be so costly that it makes them worse off at the optimum compared with the non-cooperative equilibrium (i.e. the situation prevailing without a treaty). To keep such countries convinced of the desirability of cooperating, the treaty might in addition provide for private-good transfers compensating for that cost. It is by now well known[7] that such transfers can be designed, and conceivably managed by an international agency,[8] ensuring that the condition $x_i^* + v_i(z^*) \geq \bar{x}_i + v_i(\bar{z})$ be met for each potential signatory, taken individually.

Subgroups of countries – henceforth called 'coalitions' – should also be considered, because for various reasons they may wish to act jointly instead of in cooperation with the full set N of the countries involved in the transfrontier problem. But acting in this way would mean designing treaties for themselves, involving abatement policies most likely different from $(p_1^*, \ldots, p_n^*)$ – thus suboptimal at the world level. Is this inevitable, or can it be avoided? This is exactly the point on which the controversy arises that is to be discussed presently.

Two alternative theses exist: on the one hand there is what I call the 'small stable coalitions' (SSC) thesis, according to which only small subsets of the n countries can ever emerge and sign a treaty; there is on the other hand the 'grand stable coalition' (GSC) thesis, presenting the contents of a feasible treaty which is shown to be in the interest not only of all members individually but also of all subgroups of N. These two views are developed in the next section.

3. THE TWO THESES: A SUMMARY PRESENTATION

The 'Small Stable Coalitions' (SSC) Thesis

This thesis has been formulated prominently by Carraro and Siniscalco (1993) and Barrett (1994). It is based on a concept of coalitions stability (due to d'Aspremont and Gabszewicz, 1986) borrowed from the industrial organization literature on cartels. I follow here – and limit myself to – the first authors' presentation in Carraro and Siniscalco (1995) (hereafter CS).

Let $S \subseteq N$ be a 'coalition', i.e. a set of countries that are willing to cooperate and to sign among themselves a treaty to that end. Using in this subsection the authors' notation, let $P_i(S)$ denote the utility of country i if i is a member of S, and $Q_i(S)$ denote the utility[9] of i if i is *not* a member of S.

Definition 4: The coalition $S \subset N$ is called *stable* if it satisfies the following two conditions:

(i) *internal stability*: $\forall_i \in S,\ P_i(S) \geq Q_i(S \setminus \{i\})$ and
(ii) *external stability*: $\forall_j \notin S,\ P_j(S \cup \{j\}) \leq Q_j(S)$.

For the coalition $S = N$, only internal stability applies.

For the proponents of the SSC thesis, treaties are only likely to be signed by subsets of countries that meet these two conditions. As to the likelihood of worldwide environmental treaties on worldwide pollution problems (typically climate change, where N is the set of all countries in the world), these authors are led to pessimism because of the following result (henceforth, I denote by S^* a stable coalition):

Proposition CS: *If all countries are assumed to be identical,*

(a) the existence of stable coalitions can be established;
(b) the size of stable conditions is always small, in the sense that $\forall S^*,\ |S^*| << |N|$;
(c) introducing private-good transfers between countries does not increase the size of stable coalitions.

A second result,[10] due to Botteon and Carraro (1995 – BC hereafter), mitigates the pessimism of the one just quoted. It is also based on the more realistic premise of non-identical countries. But it rests on a numerical example only:

Proposition BC: *When countries are not identical, a numerical example with five countries shows*

(a) *the existence of stable coalitions;*
(b) *that, without transfers, stable coalitions are always small;*
(c) *that private-good transfers can be found that increase the size of stable coalitions, all the way to making stable even the grand coalition N.*

As for transfers, it should be pointed out that those considered by these authors are not linked with the countries' emissions: they are all formulated as lump-sum transfers.

The 'Grand Stable Coalition' (GSC) Thesis

This thesis has been formulated and defended in the two papers Chander and Tulkens (1995, 1997) (CT hereafter; we use here mainly the model of the latter). It is based on the cooperative game-theoretic concept of the γ-core,[11] and can be summarized in the following two steps:

Assumption 'γ': If a coalition $S \subset N$ forms, the highest aggregate utility it can achieve for its members is given by the function

$$w^{\gamma}(S) = \underset{\{(x_i, p_i)_{i \in S}\}}{Max} \sum_{i \in S} [x_i + v_i(z)]$$

subject to

$$\sum_{i \in S} x_i \leq \sum_{i \in S} g_i(p_i)$$

and

$$\sum_{i \in S} p_i + z = - \sum_{j \in N \setminus S} p_j,$$

where $\forall j \in N \setminus S$, (x_j, p_j) maximizes $x_j + v_j(z)$
subject to

$$x_j \leq g_j(p_j)$$

and

$$p_j + z = -\sum_{\substack{i \in N \\ i \neq j}} p_i.$$

If coalition N forms, the highest aggregate utility it can achieve for its members is given by

$$w^{\gamma}(N) = \sum_{i \in N} \left[x_i^* + v_i(z^*) \right],$$

where x^*, z^* are values given by an international optimum $(x^*, y^*, p^*, z^*, T^*)$.

Proposition CT: *Given the vector of optimal emissions* $(p_1^*, \ldots, p_n^*)$, *private-good transfers of the form*[12]

$$T_i^* = -\left[g_i(p_i^*) - g_i(\bar{p}_i) \right] + \frac{\pi_i(z^*)}{\pi_N(z^*)} \left[\sum_{i \in N} g_i(p_i^*) - \sum_{i \in N} g_i(\bar{p}) \right], \; i \in N \qquad (1)$$

induce a feasible state $(x^*, y^*, p^*, z^*, T^*)$ *of the economic–ecological system which is such that, for every coalition* $S \subset N$,

$$\sum_{i \in S} \left[x_i^* + v_i(z^*) \right] > w^{\gamma}(S).$$

The proposition asserts that with transfers defined as in (1), the feasible state $(x^*, y^*, p^*, z^*, T^*)$ cannot be improved upon to the benefit of its members by any coalition $S \subset N$. Technically, the feasible state $(x^*, y^*, p^*, z^*, T^*)$ is a strategy that belongs to the core of a cooperative game associated with the economic–ecological system, $w^{\gamma}(S)$ being the characteristic function of that game.

The assumption yielding the function $w^{\gamma}(S)$, on which the proposition's statement rests, specifies that if S forms, its members choose the actions that are the most beneficial for themselves as a group, while the other players (countries) act to the best of their individual interests, 'playing Nash' against S and the other countries. The outcome of these behaviours is a state of the economic–ecological system that the authors call a 'Partial Agreement Nash Equilibrium with respect to S' (denoted henceforth as P.A.N.E. w.r.t.(S)). The core property of the state $(x^*, y^*, p^*, z^*, T^*)$ is thus that if a treaty is proposed to N that induces this state, no subset S of countries can hope to gain from inducing instead a P.A.N.E. w.r.t. itself. Therefore the 'grand treaty' should be signed by all, without regret.

Let me briefly recall the structure of the transfers formula (1). Each individual transfer consists of two parts: a payment *to* each country i that covers its increase in cost between the Nash equilibrium and the optimum (first squared bracket), and a payment *by* each country i of a proportion $\pi_i(z^*)/\pi_N(z^*)$ of the total of these differences across all countries (second squared bracket). In other words, each country's abatement cost is covered and each country's contribution is determined by the relative intensity of its preferences for the public-good component (z^*) of the problem. Notice that the sum of these transfers is equal to zero: they break even.[13]

Most important for our present purposes is also to note that the transfers are linked to the emissions – actually to the level of their abatement cost.

4. DIFFERENCES AND SIMILARITIES

As the preceding summary has already made clear, a theory of stable coalitions is here opposed to the theory of the core of a cooperative game. We consider here four aspects of this opposition.

On Coalitions, Coalition Formation, and the Final Outcome of the Games

Let us remind ourselves first that the theory of the core of a cooperative game, on which the GSC thesis rests, is basically *not* a theory of the formation of coalitions. Its scope is in fact more limited. It does indeed focus on arguments to support the view that only the so-called 'grand coalition'[14] of all players will form, and that the other coalitions will not form.

By contrast to this, the SSC approach claims to be able to identify some specific subsets of N for which it asserts that they will form as 'coalitions' because they are stable (in the specific SSC sense), and other subsets that will not form in this way. The justification for the assertion of stability is provided by comparing, for each conceivable subset, the payoffs of each individual player when s/he belongs to the coalition and when s/he stays out.

It thus appears that the term 'coalition' is not used in the same way by the two groups of authors. In the language of the SSC view held by the latter group, a coalition denotes a set of 'good' guys, who do cooperate among themselves,[15] and intend to sign a treaty together – while those who stay outside of the coalition are the 'bad' guys, who act in isolation. Note that, in this parlance, any coalition must comprise at least two players (here, countries): singletons are meaningless as 'coalitions'.

In the GSC (core-theoretic) way of reasoning, things are reversed: the strategy in the core (that is, the contents of a treaty for N) is supposed to be first proposed to all players; and then the term 'coalition' is used to denote people who might possibly object to it. Coalitions are thus here a set of 'bad guys', who put in question the fact of cooperating within N, and refuse to sign the grand treaty proposed to them; they instead consider doing something else – specifically, achieving what was specified above as a P.A.N.E. w.r.t.(S). Note that here a singleton *is* meaningful as a 'coalition', because the essence of a coalition is not the fact that its members cooperate, as is the case above; it is instead the fact that the coalition does (or envisages to do) something different from what is being proposed to N.

With this clarification of the vocabulary in mind, as well as of the behaviours this vocabulary is intended to describe, one can perhaps better see the central difference between the two theses, which lies in the final outcome of the transfrontier pollution game that they each envisage:

- For the SSC literature, the final outcome is a twofold situation consisting of, on the one hand, the formation of some small coalition of countries whose members do sign an abatement treaty and, on the other hand, the other countries who decline to join in signing (and enjoy a free ride from the signatories' clean-up: more on this below). Note that this outcome is in fact exactly what CT have dubbed a P.A.N.E. with respect to some coalition.
- In the GSC literature, the final outcome is a joint strategy for all players – the grand treaty, which is better for any coalition S than the P.A.N.E. this S might achieve.

In terms of the cooperative game theoretic literature on 'stable coalitions structures',[16] where a coalition structure is defined as a partition of the all-players' set, one can restate the above as follows: the SSC literature predicts an outcome with a coalition structure of the form $\{S, \{j\}_{j\in N\backslash S}\}$, where the sets $\{j\}$ are singletons, whereas the GSC literature predicts an outcome with a coalition structure of the form $\{N\}$, with no singletons. Neither of the two views under study here refers to the concept of stable coalitions structures. But it obviously applies very well to what we are dealing with.

On Free-riding and Threats

Just as with 'coalition' the expression 'free-riding' is also used with different meanings in the two strands of literature under review.

In the SSC approach, the free-riding that is dealt with is one that occurs when – in the words of its authors – 'a country lets other countries sign a cooperative agreement, and thereby enjoys a cleaner environment at no cost' (Carraro and Siniscalco, 1995, pp. 264–5). This prompts two remarks:

(i) It refers to *individual* free-riding only. Of course, one may rephrase the definition and speak of countries instead of just one. This is indeed the case with the SSC final outcome I have just recalled. But the set of such free-riders then amounts to a collection of singletons, not a set of cooperating players.

(ii) Suppose a (for example upstream) country is a major polluter, but does not care at all for the quality of the (downstream) environment, for objective reasons. As it pollutes a lot, it should be brought into the treaty, since its actions are determinant ones for achieving a full international optimum. If it stays out nevertheless, is it to be considered as a free-rider? In fact, the above definition of free-riding does not apply very well to such a case.

Turning to the GSC view, I immediately see two elements emanating from the core concept that are relevant to free-riding:

(i) Free-riding is considered for *any* subset S of N, that is, for singletons but also for larger subsets of N. We have thus explicitly the possibility of *coalitional* free-riding.

(ii) It is supposed that free-riders do cooperate among themselves: they indeed are assumed to achieve $w^{\gamma}(S)$, as defined in Assumption γ.

Much more importantly, however, the GSC view adds another ingredient in describing free-riding behaviour, namely a reaction of the other, non-free-riding countries. This reaction is not intended to punish the free-riders in an irrational way; it is simply not to form as a coalition, and to just play Nash against the free-riding coalition S. This is a threat element that I like to call an individually reasonable threat.

Threats against free-riders[17] are absent from the SSC analysis; but the constructive results yielded by the GSC analysis, using some form of threat, make one wonder whether this is not precisely an important source of the difficulty, for the former, in finding grounds for cooperative agreements.

Characteristic Functions: A Common Tool for Further Analysis

We have observed above that the final outcome of the SSC approach is nothing else than a P.A.N.E. w.r.t.(S) where some S is found to be stable. On the other hand, that same concept is used by the GSC approach to formulate the characteristic function $w^{\gamma}(S)$ whereby a coalitionally stable strategy is claimed to be found for N.

This rapprochement suggests that while the SSC thesis does not use the tool of a characteristic function, one could nevertheless ask whether there is not some characteristic function underlying, or hidden within, the SSC approach. I want to argue here that this is indeed the case, after having made two preliminary remarks on the characteristic function $w^{\gamma}(S)$.

Let me observe, first, that with the characteristic function $w^{\gamma}(S)$, in the special case where $S = \{i\}$ is a singleton, the resulting P.A.N.E. w.r.t. $(\{i\})$ is nothing else than the Nash equilibrium of the problem. Any individual free-riding, in the GCS framework of thought, entails absence of any cooperation at all. This is the extreme form of the threat I described above.

Second, there is also something to be learned from considering, still with the characteristic function $w^{\gamma}(S)$, the other extreme case where $S = N\backslash\{i\}$, and the final outcome is the P.A.N.E. w.r.t.$(N\backslash\{i\})$. Here, $N\backslash\{i\}$ are cooperating (thus, they are 'coalitional' free-riders), and $\{i\}$ is left alone. Compared with the previous case, things are reversed. The 'free-rider' expression is perhaps not too appropriate a vocabulary any more, since the outcome may be more naturally seen as what occurs when the full players' group N throws out the singleton $\{i\}$.

What do we learn from considering this case? Essentially that the core strategy for N is to be understood as one that deters $N\backslash\{i\}$ to act that way. This is relevant for the case mentioned above, namely when i is a strong polluter, careless of the environment, and neglecting to cooperate with N: the core strategy is one such that for the members of $N\backslash\{i\}$, it is not in their interest to leave $\{i\}$ out.

My main point in this subsection is a different one, however. In the definition of the characteristic function $w^{\gamma}(S)$, it is assumed that, given S, the players not in S play Nash against this coalition, and $w^{\gamma}(S)$ then denotes the payoff for the members of S, given that assumption.

Now, why not change this assumption, and consider what the SSC literature denotes as the magnitude $Q_i(S)$, that is, the payoff of player i when he is not a member of S, *and S is formed.*[18] Let us, in particular, consider this magnitude for $S = N \setminus i$, that is, $Q_i(N \setminus i)$. Using now the variables of the underlying economic–ecological model, let us exhibit the strategies of all players that induce such a payoff. In the notation used earlier, we have:

$$Q_i(N \setminus i) \equiv \mathrm{Max}\ u_i = x_i + v_i(z)$$

$$\text{s.t. } x_i \leq g_i(p_i)$$

$$p_i + z = -\sum_{j \neq i} p_j$$

where the vector $(p_j)_{j \in N\backslash i}$ maximizes

$$\sum_{j \in N \setminus i} u_j = \sum_{j \in N \setminus i} \left[x_j + v_j(z)\right]$$

$$\text{s.t.} \sum_{j \in N \setminus i} x_j \leq \sum_{j \in N \setminus i} g_j(p_j)$$

$$\sum_{j \in N \setminus i} p_j + z = -p_i.$$

To harmonize notation, let me now substitute[19] $w^{\delta}(\{i\})$ for the value of $Q_i(N \setminus i)$ so defined. Let me further define this value $w^{\delta}(\{i\})$ for all singletons of N, and write for N itself $w^{\delta}(N) = w^{\gamma}(N)$ as defined in Assumption γ.

I thus define a function $w^{\delta}(\cdot)$ that associates with all singletons of N, and N itself, a real number. In cooperative game-theoretic parlance, this is of the nature of a characteristic function, with the peculiarity that its domain is restricted to only some subsets of N. Nevertheless, we have a cooperative game, defined by the pair $[N, w^{\delta}(\cdot)]$.

If for this game a core imputation exists, then N is a stable coalition *in the SSC sense*, and we have a reconciliation of the two theses. If the core is empty, then N is not a stable coalition in that sense, in spite of Proposition CT: the two concepts cannot be reconciled, in general.

This is what Proposition CS establishes, using the case of identical players: the core of the game $[N, w^{\delta}(\cdot)]$ is thus empty, in general. However, Botteon and Carraro (1995) showed, with an example, that with non-identical players and transfers, the core of that game may not be empty: reconciliation is thus not a hopeless task. It only remains to find out how strong, and realistic, general conditions can be under which non-emptiness holds. I have not done that work, but I am convinced that it would be worth doing, if only because the economic–ecological world we are dealing with is essentially and immensely diversified, and transfer of resources across countries is evidently a tool of international economic policy.

With the construct just presented, I have attempted to reformulate the stable coalitions theory in terms of the theory of cooperative games. The scope of that attempt is of course limited to the issue of the stability of the grand coalition *vis-à-vis individual*[20] free-riding, and its interest essentially rests in delineating the conditions of reconciliation[21] between the two theories.

On Transfers and 'Side Payments'

A final dissimilarity lies in the formulation of transfers. As pointed out in Section 3, they are of the lump-sum form in the SSC models, whereas in Formula (1) of the GSC approach they appear as linked with the amounts of emissions abatement. While in the former case they are just 'side payments' between countries, they can be given in the latter case an interpretation in terms of a formula[22] for sharing, between the countries, the aggregate abatement costs.

Introducing this second kind of transfer in the characteristic function apparatus I have just outlined would definitely be relevant. While such transfers would not change the negative result obtained by CS with identical countries, they might reinforce the positive result of Botteon and Carraro with non-identical countries.

5. CONCLUSION

What is essentially at stake in this controversy is the stability of the grand coalition: are all countries likely to sign treaties in matters of worldwide transfrontier pollution problems? The above comparative exercise suggests an answer in the form of a further question: what kind of stability does one have in mind: (1) a *passive* stability with respect to *singletons only,* with 'passive' meaning stability without threat against defecting singletons? - this is the SSC view; or (2) an *active* stability with respect to *all conceivable coalitions,* with 'active' meaning stability with the threat of playing Nash against defecting coalitions? - this is the GSC view.

From a positive economics point of view, both concepts are defensible, and it remains to the analyst to find out which one is more often observed, and therefore more realistic. From a normative point of view, in which I would include the discourse of policy advisers, I cannot help thinking that the active stability perspective has stronger merits for two reasons: it embodies the reality of threats in a richer way; and it has shown that it lends itself to formulating explicit emissions and transfers policies that are both implementable and computable.

Yet there is, of course, a long way, a very long way indeed, between what our modest models allow us to assert, and the immensely complex reality we are facing. But I cannot help being happy with theoretical thinking that gives some ground for optimism, because in this way it becomes possible that our intellectual and scientific activity contributes positively to the endeavours of negotiators and decision makers who are in charge of those matters. When theory can help them in a constructive way, I submit we do our job best.

NOTES

1. Invited keynote speech at the Sixth Meeting of the European Association of Environmental and Resource Economists, Umea, Sweden, 22 June 1995. The research reported here was part of the activities carried out under support of the European Union, within the project 'Environmental Policy, International Agreements and International Trade' coordinated by Professor Alistair Ulph, as well as within the HCM Network 'Designing economic policy for

management of natural resources and the environment' (CHRX CT93-0228) coordinated by Professor Anastasios Xepapadeas. It originated much earlier, however, during stimulating and valuable discussions with Carlo Carraro and Domenico Siniscalco that started at the Cambridge 1991 meeting of the European Economic Association and were pursued at several other occasions, in particular at the Harvard 1994 Congress of the International Institute of Public Finance and at the CORE–FEEM workshop 'Coalitions in environmental games and other economic applications' held at the Center for Operations Research and Econometrics (CORE), Louvain-la-Neuve in March 1995. For the completion of the chapter the author also benefited from the comments of Claude d'Aspremont and the editors of this volume, as well as from the hospitality and support of the Fondazione ENI Enrico Mattei (FEEM), which he gratefully acknowledges.

2. More detailed presentations, with some discussion of the main assumptions can be found in Section 2 of Chander and Tulkens (1992).
3. The absence of a subscript attached to this variable reflects its public-good character; and with the convention of measuring the ambient characteristic in non-positive amounts, our assumption $\partial u_i/\partial z \geq 0$ implies that z is felt by the consumers as a public bad. Notice also that z is treated in this chapter as a flow only. Extensions to stock pollutants have been made recently for the GSC thesis (expounded below) in Germain, Toint and Tulkens (1996) as well as in Germain, Tulkens and de Zeeuw (1996). Another limitation of all models discussed here is that they deal with scalar-measured pollutants only.
4. Labour, capital and the other inputs are taken as constant and subsumed in the functional symbol g.
5. With some abuse of language, these transfers will often be called 'financial' in what follows.
6. Close analysis of the economic model reveals that, just as in reality, there are many optima in general – optima that may differ either in terms of the emissions vector $(p_1^*, \ldots, p_n^*)$, or in terms of the consumption levels $(x_1^*, \ldots, x_n^*)$, or both. The quasi-linearity assumption simplifies the reasoning in this respect because it implies that the emissions vector $(p_1^*, \ldots, p_n^*)$ is the same at all international optima (for a proof, see Proposition 1 in Chander and Tulkens, 1995b).
7. As amply elaborated upon in Chander and Tulkens (1991 and 1992). For a more sceptical review, see, however, Folmer, Mouche and Ragland (1993, pp. 314–15).
8. As suggested in Tulkens (1979, p. 206).
9. These utilities could also be written in the notation of the previous section, but I shall turn to that later.
10. I shall leave aside the otherwise interesting results of CS concerning the implications of possible *commitments* to cooperate, because the issue here is only the explanation of cooperation.
11. As distinct from those of α- and β-cores. The Greek letters used refer to alternative specifications of the assumption just about to be stated. The references cited contain a discussion of these alternative assumptions.
12. Notation: $\pi_i = dv_i/dz$ and $\pi_N = \Sigma_{i \in N} \pi_i$
13. Several further comments and properties are given in Chander and Tulkens (1995, pp. 289–91).
14. As will appear below, it may be expositionally convenient to keep the term 'coalition' for *proper* subsets of players only and avoid using it for denoting the full players' set. As an additional justification for this terminological convention, one may remember that a coalition is usually conceived of as a group opposing itself *against* some other people. Clearly there are no such other people when the 'coalition' is the full players' set.
15. In both theories, 'to cooperate' means the same thing, namely: for any set of players whose cardinal number is at least two, to do together something different from what each of the cooperating players would do alone.
16. A concept used and studied by Aumann and Drèze (1974), as well as Hart and Kurz (1983).
17. Threats are of course not to be confused with the 'reaction functions' analysed with much detail in section 3.2 of CS. On reaction functions, however, it is interesting to note that while stability in the SSC sense is shown by CS to be weakened by non-orthogonal functions, the P.A.N.E.(S) on which stability of N in the GSC sense is established do imply non-orthogonal reaction functions (see assertion (*iii*) in Proposition 4 of Chander and Tulkens, 1997).

18. This proviso was not explicitly mentioned earlier; but it is unquestionably present, albeit implicitly, in the SSC theory.
19. The superscript δ is used to point out to the once more different assumption made here on the behaviour of players that do *not* belong to the coalition S under consideration. Writing this δ case in a more explicit way as $w^{\delta}(\{i\} \mid N \setminus \{i\})$, one could also imagine still further cases suggested by the expression $w^{\varepsilon}(\{i\} \mid S, \{j\}_{j\neq i, j\notin S})$ Examining these is beyond the scope of the present discussion.
20. Thus, not coalitional free-riding.
21. When the core is empty, one may consider as a substitute the nucleolus.
22. The details of this are given in section 6 of Chander and Tulkens (1995).

REFERENCES

Aumann, R.J. and J. Drèze (1974), 'Cooperative games with coalition structures', *International Journal of Game Theory*, **3**, 217–37.

Barrett, S. (1994), 'Self-enforcing international environmental agreements', *Oxford Economic Papers*, **46**, 878–94.

Botteon, M. and C. Carraro (1995), 'Burden sharing and coalition stability in environmental negotiations with asymmetric countries', paper presented at the 2nd meeting of the HCM network, 'Designing economic policy for management of natural resources and the environment', Venice, May 1995. FEEM Nota di Lavoro no. 78.95 (November).

Carraro, C. and D. Siniscalco (1993), 'Strategies for the international protection of the environment', *Journal of Public Economics*, **52**, 309–28.

Carraro, C. and D. Siniscalco (1995), 'International coordination of environmental policies and stability of global environmental agreements', chap. 13 in L. Bovenberg and S. Cnossen (eds), *Public economics and the environment in an imperfect world*, Boston, London, Dordrecht: Kluwer Academic Publishers.

Chander, P. and H. Tulkens (1991), 'Strategically Stable Cost Sharing in an Economic–Ecological Negotiation Process', CORE Discussion Paper no. 9135, June 1991 (revised October 1992); to appear in K.G. Mäler (ed.), *International Environmental Problems*, a volume sponsored by the European Science Foundation, Boston and Dordrecht: Kluwer.

Chander, P. and H. Tulkens (1992), 'Theoretical foundations of negotiations and cost sharing in transfrontier pollution problems', *European Economic Review*, **36** (2/3), 288–99.

Chander, P. and H. Tulkens (1995), 'A core-theoretic solution for the design of cooperative agreements on transfrontier pollution', *International Tax and Public Finance*, 2 (2), 279–94.

Chander, P. and H. Tulkens (1997), 'The Core Of An Economy With Multilateral Environmental Externalities', *International Journal of Game Theory*, **26**, 379–401.

d'Aspremont, C. and J. Jaskold Gabszewicz (1986), 'On the stability of collusion', chapter 8 (pp. 243–64) of J.E. Stiglitz and G.F. Mathewson (eds), *New developments in the analysis of market structure*, Cambridge, MA: MIT Press.

Folmer, H., P. van Mouche and S. Ragland (1993), 'Interconnected games and international environmental problems', *Environmental and Resource Economics*, **3**(4), 313–36.

Germain, M., P. Toint and H. Tulkens (1996), 'Financial transfers to ensure cooperative international optimality in stock pollutant abatement', CORE Discussion Paper no. 9701, Center for Operations Research and Econometrics (CORE), Université

Catholique de Louvain, Louvain la Neuve; to appear in F. Duchin, S. Faucheux, J. Gowdy and I. Nicolaï (eds), *Sustainability and Firms: Technological Change and the Changing Regulatory Environment*, Cheltenham: Edward Elgar.

Germain, M., H. Tulkens and A. de Zeeuw (1996), 'Stabilité stratégique et pollution-stock transnationale: le cas linéaire', IRES Discussion Paper no. 9701, Université Catholique de Louvain, Louvain-la-Neuve, forthcoming in *Revue Economique*, 1998.

Hart, S. and M. Kurz (1983), 'Endogenous formation of coalitions', *Econometrica*, **51** (4), 1047–65.

Tulkens, H. (1979), 'An Economic Model of International Negotiations Relating to Transfrontier Pollution', chapter 16 (pp. 199–212) in K. Krippendorff (ed.), *Communication and Control in Society*, New York: Gordon and Breach Science Publishers.