DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

A CORE–THEORETIC SOLUTION FOR THE DESIGN OF COOPERATIVE AGREEMENTS ON TRANSFRONTIER POLLUTION

Parkash Chander
Indian Statistical Institute and California Institute of Technology

Henry Tulkens
CORE, Université Catholique de Louvain

SOCIAL SCIENCE WORKING PAPER 897

August 1994

A Core-Theoretic Solution for the Design of Cooperative Agreements on Transfrontier Pollution

Parkash Chander Henry Tulkens

Abstract

The paper highlights the relevance of the game theoretic concept of the core of a cooperative game for the design of international treaties on transfrontier pollution. Specifically, a formula is offered for allocating abatement costs between the countries involved for which the justification is of core-theoretic nature. The analysis emphasizes the strategic role of monetary transfers among the countries.

A Core-Theoretic Solution for the Design of Cooperative Agreements on Transfrontier Pollution*

Parkash Chander Henry Tulkens

1 Introduction

The purpose of this paper is to highlight the relevance of the game theoretic concept of the *core* of a cooperative game for the design of international treaties on transfrontier pollution. Specifically, a formula is offered (in Section 5) for allocating abatement costs between the countries involved for which the justification is of core-theoretic nature.

We develop our arguments in the framework of the simplest model traditionally used for the economic analysis of such agreements (as e.g. in MÄLER 1989-93, HOEL 1992, BARRETT 1992, CARRARO and SINISCALCO 1993, d'ASPREMONT and GERARD-VARET 1992.) Our claims proceed, however, from results presented with full technical details in a companion theoretical paper (CHANDER and TULKENS 1994, hereafter referred to as C&T94), where use is made of a more general model expressed in terms of an Arrow-Debreu economy, initially formulated in TULKENS 1979 as well as in CHANDER and TULKENS 1992. In the latter paper, a cost sharing formula in the same spirit was presented, valid only for marginal adjustments in abatement. Here, the formula is extended to the global game.

*Thanks are due to Karl Göran Mäler for numerous fruitful discussions and his hospitality at the Beijer Institute for Ecological Economics, Royal Swedish Academy of Sciences, Stockholm, during May-June 1993. The first author is also grateful to California Institute of Technology for providing a stimulating environment for the completion of this work. This research is part of the Commission of the European Communities (DG XII) "Environmental Policy, International Agreements and International Trade," administered by Alistair Ulph through CEPR, London.

2 Transfrontier Pollution: The Economic Model and Its Associated Games

2.1 The Basic Economic Model

Currently, much of the economic analysis of international agreements on transfrontier pollution is based on a model whose components are the following:

(i) $N = \{i | i = 1, \ldots, n\}$, the set of countries concerned by the analysis, the countries being n in number, each indexed by i;

(ii) For each country i:

- **(a)** Quantities $E_i \geq 0$ of pollutants *emitted* by the economic agents of country i, per unit of time. E_i is a scalar if the analysis bears on one pollutant only, as will be the case in the present paper.[1]
- **(b)** Quantities $Q_i \geq 0$ of *ambient* pollutants present in country i's environment, per unit of time. As for the emissions, Q_i is a scalar if only one pollutant is considered; in an analysis with several pollutants it would be a vector.
- **(c)** A *transfer function*

$$Q_i = F_i(E), \tag{1}$$

  where $E = (E_1, \ldots, E_n)$, that describes the physical, chemical, and/or biological processes whereby the amounts E of pollutant(s) emitted in all countries get transformed into the quantities Q_i of ambient pollutants present in country i. This function is assumed to be nondecreasing in each of its arguments. The fact that Q_i is formulated here as being dependent on *current* emissions only restricts the analysis to *flow* pollutants, as opposed to *stock* pollutants.
- **(d)** An *abatement cost function*, $C_i(E_i)$, expressing the costs (monetary and possibly nonmonetary) incurred by the polluting agents of country i when their aggregate emissions are restricted to the amount E_i. This function is assumed to be decreasing ($C_i' < 0$), a property reflecting the natural assumption that reducing emissions (i.e., abating) is costly.
- **(e)** A *damage cost function* $D_i(Q_i)$, expressing the costs (monetary as well as nonmonetary) incurred by the economic agents of country i as a result of the ambient pollutants Q_i they are exposed to. This function is assumed to be nondecreasing ($D_i' \geq 0$).
- **(f)** The *total of abatement and damage costs*

$$\begin{aligned} J_i(E) &\equiv C_i(E_i) + D_i(Q_i) \\ &= C_i(E_i) + D_i[F_i(E)], \end{aligned} \tag{2}$$

[1] It would be an m-dimensional vector if the quantities of m different pollutants were considered.

incurred by the country as a result of the joint pollutant emissions E of all countries. Notice that the variables $E_j, j \neq i$, that appear as arguments of the function J_i are of the nature of an externality, exerted on country i by each one of the countries j.

(iii) Finally, the description of this international economy with pollution—henceforth summarily designated by the pair $[N, (C_i, D_i, F_i)_{i \in N}]$—is completed by identifying a vector $E^* = (E_1^*, \ldots, E_n^*)$ of *optimal joint emissions* by the n countries, optimality being taken in the sense of minimizing the sum over all countries of both abatement and damage costs. E^* is thus the solution of the optimization problem.

$$\min_{\{E_1, \ldots, E_n\}} J(E), \tag{3}$$

where $J(E) \equiv \sum_{i \in N} J_i(E)$.

2.2 The Associated Games

Because of the nonexistence of a supranational authority endowed with sufficient coercive power to impose any emissions policy on the countries, the optimum just defined is only likely to be implemented on a *voluntary* basis, that is, in terms of joint actions that suit the interests of each one of the countries involved.

This argument, which is by now classical, motivates the recourse to game theoretic concepts for finding out whether it is at all reasonable to expect that the optimum be achieved voluntarily by the parties. Indeed, the formulation of alternative games associated with the economic model introduces behavioral assumptions on the basis of which voluntary actions can be characterized.

In this respect, the distinction offered by classical game theory between noncooperative and cooperative games is particularly relevant. This paper builds explicitly on such distinction (i) by characterizing as equilibria of a *noncooperative game* associated with the above economic model, national emission policies that only satisfy the objectives of each country; and (ii) by identifying with some solution concept for *cooperative games* (also associated with the economic model) policies that reflect actions taken in a coordinated way by either all the parties, or subsets of them.

Formally,

(i) **A noncooperative game**, defined by its players set $N = \{i | i = 1, \ldots, n\}$, by the sets $T_i, i = 1, \ldots, n$ of strategies accessible to each of the players i, and by the payoffs u_i that the latter achieve—and henceforth denoted by the triplet $[N, (T_i)_{i \in N}, (u_i)_{i \in N}]$—is associated with the economic model $[N, (C_i, D_i, F_i)_{i \in N}]$ by identifying the players set with

the set of countries, by defining the strategy set of each country as $T_i = \{E_i | E_i \geq 0\}$ (with possibly an upper bound E_i^o to be defined below), and by defining each player's payoff u_i as the value $-J_i$ of the function specified in (2) above.

(ii) **A cooperative game** (in characteristic function form and with transferable utility), defined by its players set $N = \{i | i = 1, \ldots, n\}$ and the function $w(S)$ that associates with every subset S of N a number called the worth[2] of S—and henceforth denoted by the pair $[N, w]$—is similarly associated with the economic model $[N, (C_i, D_i, F_i)_{i \in N}]$ by identifying again the player set with the set of countries, and by defining the characteristic function as[3]

$$w(S) = \min_{(E_i)_{i \in S}} \sum_{i \in S} [C_i(E_i) + D_i(Q_i)]. \tag{4}$$

Thus the worth of each coalition S is determined by some strategy vector $(E_i)_{i \in S}$ adopted by the members of the coalition. However, remembering (1), one notices that when $S \neq N$, this worth also depends upon the strategies $(E_j)_{j \notin S}$ adopted by the countries which are not members of S. As those have been left unspecified in our formulation of the function (4), we shall have to return to this issue below when we deal in more detail with the cooperative game.

2.3 Assumptions on the Economic and Ecological Components of the Model

Precise results on voluntary behavior in this economic model can be obtained when some further assumptions are introduced on its components. Those we shall use—most of which are standard, but would deserve critical discussion—are the following:

Assumption 1: For every (decreasing) abatement cost function $C_i(E_i), i = 1, \ldots, n$, there exists $E_i^o > 0$ such that

$$C_i'(E_i) \begin{cases} = -\infty, & \text{if} \quad E_i = 0, \\ < 0 & \text{if} \quad E_i < E_i^o \\ = 0, & \text{if} \quad E_i \geq E_i^o. \end{cases} \tag{5}$$

Assumption 2: $\forall i$, the function $C_i(E_i)$ is strictly convex (i.e., $C_i'' > 0$) over the range $]\, 0, E_i^o \,[$.

[2]or the payoff to coalition S.

[3]In standard game theoretic models, payoffs, whether individual as in u_i or for coalitions as in $w(S)$, are usually supposed to be maximized by the players. As the economic model used here associates costs only with its agents, maximizing payoffs for them amounts to minimizing costs. The Arrow-Debreu type of model used in C&T94 allows for payoffs to be directly defined on the utilities of the economic agents, more in line with usual practice.

Assumption 3: $\forall i$, the transfer function $Q_i = F_i(E)$ is of the linear additive form

$$Q_i = \sum_{j=1}^{n} E_j. \tag{6}$$

Notice that this assumption implies that $Q_i = Q_j \ \ \forall \ \ i, j \in N$, thus making the ambient quantities of pollutant to have the characteristics of an international public good (actually, of a public "bad") for the countries involved.[4] In view of Assumption 3, we shall often write, with some notational inconsistency, $D_i(Q_i)$ as $D_i(E)$.

Assumption 4: $\forall\ i$, the (nondecreasing) damage cost function $D_i(Q_i)$ is convex ($D_i'' \geq 0$); it is strictly increasing ($D_i' > 0$) for at least some i.

Together, Assumptions 2 to 4 imply that $\forall\ i$, the total cost function $J_i(E)$ is convex. One can then prove, as in C&T94 (Section 3):

Proposition 1: Under the Assumptions 1-4, the optimal joint emissions vector E^* is unique and in the range $]\ 0, E_i^o\ [\ \ \forall i$.

The optimum so defined is usefully characterized by the well-known first order conditions

$$\sum_{j=1}^{n} D_j'(E^*) + C_i'(E_i^*) = 0, \quad i = 1, \ldots, n. \tag{8}$$

3 The Noncooperative Game and Its Nash Equilibria

3.1 The Nash Equilibrium

A first form of voluntary behavior in our economic model is the one described by the familiar Nash equilibrium concept of the associated noncooperative game, namely:

Definition 1: For the noncooperative game $[N, (T_i)_{i\in N}, (u_i)_{i\in N}]$, a Nash equilibrium is a joint strategy choice $\bar{E} = (\bar{E}_1, \ldots, \bar{E}_n)$ such that $\forall\ i, \bar{E}_i$ minimizes $J_i(E)$, where for each

[4] In MÄLER's 1989-93 acid rain game, where the linear transfer function is of the form

$$Q_i = \sum_{j=1}^{n} a_{ji} E_j, \tag{7}$$

with $0 \leq a_{ji} \leq 1$ and $\sum_{i\in N} a_{ji} = 1$, the externalities are directional and do not have the public good property.

$j \in N, j \neq i, E_j = \bar{E}_j$.

Existence of this equilibrium follows from standard theorems (see e.g. FRIEDMAN 1990). It is characterized by the first order conditions

$$D_i'(\bar{E}) + C_i'(\bar{E}_i) = 0, \quad i = 1, \ldots, n. \tag{9}$$

As these differ from the conditions (8) whereby the optimum was characterized, a Nash equilibrium is not an optimum for the economy, revealing thus that this form of voluntary behavior is incompatible with international optimality.

Notice that the conditions (9) imply

$$\forall i, \bar{E}_i \begin{cases} < E_i^o & \text{if } D_i' > 0, \\ = E_i^o & \text{if } D_i' = 0. \end{cases} \tag{10}$$

Furthermore, if $D_i(Q_i)$ is linear, $\bar{E}_i$ is a *dominant strategy* for i since the first term of (9) is independent of Q_i and of the vector E in that case.

A final property, formally important for our purposes below, is the *uniqueness* of the Nash equilibrium vector $\bar{E}$ in this game, as shown in Proposition 2 of C&T94.

3.2 Strong and Coalition-Proof Equilibria

MÄLER 1989-93 has considered stronger concepts of voluntary behavior in the framework of noncooperative games, namely the "strong Nash equilibrium" and the "coalition-proof Nash equilibrium." However, for the game $[N, (T_i)_{i \in N}, (u_i)_{i \in N}]$, it can be shown that there exists no strong Nash equilibrium, and no coalition-proof Nash equilibrium is optimal, if there exists one at all. They are therefore of little use in our enquiry.

3.3 The Partial Agreement Nash Equilibrium With Respect to a Coalition

Another aspect of the noncooperative behavior may be considered, namely the one adopted by the players outside a coalition, when a coalition forms. This is described by the following concept.

Definition 2: Given some coalition $S \subset N$, **a partial agreement equilibrium with respect to** S in the game $[N, (T_i)_{i \in N}, (u_i)_{i \in N}]$, is a joint strategy $\tilde{E}$ such that

(i) $(\tilde{E}_i)_{i\in S}$ minimizes $\sum_{i\in S} J_i(E)$, where for every $j \in N, j \notin S, E_j = \tilde{E}_j$ as defined in (ii), and

(ii) $\forall j \in N\backslash S, \tilde{E}_j$ minimizes $J_j(E)$, where for every $i \in S, E_i = \tilde{E}_i$ as defined in (i).

In C&T94 (Section 3.3), we prove:

Proposition 2. For any proper coalition $S \subset N$ in the game $[N, (T_i)_{i\in N}, (u_i)_{i\in N}]$,

(i) there exists a partial agreement equilibrium with respect to S;

(ii) the vector of individual emission levels at such an equilibrium is unique;

(iii) the individual emissions of the players outside the coalition are not lower than those at a Nash equilibrium;

(iv) the total emissions level is not higher than at a Nash equilibrium.

The equilibrium so defined is also characterized by the first order conditions

$$\begin{aligned} \sum_{j\in S} D'_j(\tilde{E}) + C'_i(\tilde{E}_i) &= 0, \quad i \in S \\ D'_j(\tilde{E}) + C'_j(\tilde{E}_j) &= 0, \quad j \in N\backslash S. \end{aligned} \tag{11}$$

4 The Cooperative Game: Imputations, The Core and Alternative Characteristic Functions

4.1 Imputations and the Core

Turning now to the cooperative part of our analysis, we first recall some terminology. For a cooperative game $[N, w]$ in general, an *imputation* is a vector $y = (y_1, \ldots, y_n)$ such that $\sum_{i\in N} y_i = w(N)$. Recall that for the game associated with our economic model, the worth $w(N)$ of the grand coalition, as defined in (4), is a total cost: more precisely, it is the minimum of the aggregate total abatement and damage cost over all countries. Here, an imputation is thus a way to share among all players the amount of this cost. In this setting, an imputation y is said to *belong to the core* of the game if it satisfies the conditions;

$$\sum_{i\in S} y_i \leq w(S), \ \ \forall \ S \subset N. \tag{12}$$

The core of our cooperative game is thus the set of imputations having the property that to every conceivable coalition they offer to bear a share of the aggregate cost $w(N)$ lower than the cost $w(S)$, it would bear by itself.

To study the core of this game, we therefore shall consider in more detail, in the next two subsections, what its imputations are, as well as how its characteristic function is precisely defined.

4.2 Imputations in the Game and Monetary Transfers in the Economic Model

As was noted in Proposition 1, the minimum aggregate cost is determined by a unique joint strategy vector $E^* = (E_1^*, \ldots, E_n^*)$, yielding $J_i(E^*)$ for each i and of course $\sum_{i \in N} J_i(E^*) = w(N)$. The vector $J(E^*) = [J_1(E^*), \ldots, J_n(E^*)]$ is thus an imputation where each country bears itself the abatement and damage costs that E^* entails for it.

Other imputations, associated with the same optimal joint strategy, can be conceived of, however, if monetary transfers between countries are introduced. Let us denote such transfers by $P_i (> 0$ if the transfer is paid by $i, < 0$ if it is received by it). Then imputations in the cooperative game associated with our economy can be written as vectors $y^p = (y_1^p, \ldots, y_n^p)$ defined by

$$y_i^p = J_i(E^*) + P_i, \quad i = 1, \ldots, n,$$

and the condition

$$\sum_{i \in N} P_i = 0.$$

With the condition just stated, we have indeed that $\sum_{i \in N} y_i^p = w(N)$.

It was observed at the end of Subsection 2.2 that for arguments $S \neq N$ of the characteristic function associated with our economic model, the function involves variables that represent strategic choices made by players who are *not* members of S. Because of this feature—a typical one when a cooperative game is associated with economies with externalities, such as ours—the characteristic function (4) should specify explicitly what the actions are both of the members of S and of the other players. To this effect, we shall consider the following two alternatives:

(i) The cooperative game $[N, w^\alpha]$, defined by the characteristic function of the form

$$w^\alpha(S) = \min_{(E_i)_{i \in S}} \sum_{i \in S} J_i(E) \text{ where, if } S \neq N, E_j = E_j^o \; \forall \, j \in N \backslash S.$$

(Recall that E_j^o was defined in Assumption 1).

This function reflects the assumption that when a coalition forms, its worth is what it gets when the players outside the coalition choose the strategy which is worst for it—i.e., pollute up to E_j^o in our model.

This form has been often used in economic models with beneficial externalities and/or public goods production (see, e.g. FOLEY 1970, SCARF 1971, and recently CHANDER 1993) where it is a natural one because the "worst" strategy of nonmembers of S is simply no action in such cases.

With detrimental externalities as we have here, it is less natural to assume such an attitude: why should the nonmembers of S act in this way? And for the members of S, why should they necessarily expect the worst and behave in a minimax way? The first of these questions is also raised by MÄLER 1989-1993 in his discussion of cooperative games of transfrontier pollution, all the more rightly so that he does not assume in the economic model an upper bound such as our E_j^o for the individual emissions. The worst then becomes infinite amounts of emissions, which is hardly credible.

While Mäler concludes by dismissing the tool of the characteristic function, and as a consequence the core concept which is built on it[5], we choose to propose instead to consider the following alternative:

(ii) The cooperative game $[N, w^\gamma]$, defined by the characteristic function of the form

$$w^\gamma(S) = \min_{(E_i)_{i\in S}} J_i(E) \text{ where, if } S \neq N, E_j = \tilde{E}_j \forall j \in N\backslash S.$$

(Recall that $\tilde{E}_j$ was defined in part (ii) of the definition of a partial agreement equilibrium with respect to a coalition).

The function w^γ is to be called the "partial agreement" characteristic function. We assume here that when S forms, the other players break-up into singletons, and act noncooperatively so as to reach an equilibrium in their best individual interest, given S.

It is thus not assumed that they do the worst; nor is it assumed, as in the concepts of strong and coalition-proof equilibria, that they do not react[6] to the actions of S.

[5]Another argument made is that with no bounds on the behavior of players not in S, the worth of coalitions different from N can be reduced to zero, which renders them powerless. The core then becomes equal to the set of imputations, and the concept brings no more information than optimality.

[6]CARRARO and SINISCALCO 1993, in a model with identical agents, assume instead that when S forms and achieves the aggregate payoff $w(S)$, if some $i \in S$ leaves S, the coalition $S\backslash\{i\}$ remains formed. (They show that then, it may be better for i to leave S; and as this advantage grows with the size of coalitions, they conclude that only small coalitions can prevail, and N will never form).

In view of property (10) and of Proposition 2, one has that for each $S, w^\gamma(S) \geq w^\alpha(S)$. This implies that the core of the game $[N, w^\gamma]$, i.e., the "γ-core" is, if nonempty, contained in the "α-core", and possibly smaller.

5 An Imputation in the "γ-Core" of the Cooperative Game

As it is well known that many cooperative games may have an empty core, the concept of a γ-core imputation is only useful if we can establish its existence, at least for the cooperative game $[N, w^\gamma]$ that we have associated with the economic model. As far as the α-core is concerned, it was shown to be nonempty in games with externalities by SCARF 1971 as well as in the version given by LAFFONT 1977 (p. 102) of the SHAPLEY and SHUBIK 1989 well-known "garbage game"[7].

We proceed in this section in a constructive way, that is, by exhibiting an imputation for which we show that it has the property of belonging to the γ-core. Economic interpretations are given in the next section.

Our result is not fully general, though, as we obtain it only under two alternative additional assumptions: either linearity of the damage cost functions D_i, or identical abatement cost functions C_i for all countries i. We limit ourselves here to the first case, and refer the reader to C&T94 for the second one.

Theorem. Let $E^* = (E_1^*, \ldots, E_n^*)$ be the (unique) optimal joint emissions policy. Under Assumptions 1-3 and the linearity of all the damage cost functions D_i, the imputation $y^* = (y_1^*, \ldots, y_n^*)$ defined by

$$y_i^* = J_i(E^*) + P_i^*, \quad i = 1, \ldots, n,$$

where

$$P_i^* = -[C_i(E_i^*) - C_i(\bar{E}_i)] + \frac{D_i'}{D_N'}\left[\sum_{i \in N} C_i(E_i^*) - \sum_{i \in N} C_i(\bar{E}_i)\right] \tag{13}$$

and

$$D_N' = \sum_{i \in N} D_i',$$

[7]Laffont also shows that for the garbage game, the emptiness claimed by Shapley and Shubik applies in fact to the β-core. For a game like ours, the α-core and the β-core coincide.

belongs to the core of the game $[N, w^\gamma]$.

Proof: (a) It is easily verified that y^* is an imputation, i.e., that

$$\sum_{i \in N} y_i^* = w^\gamma(N) = \sum_{i \in N} [C_i(E_i^*) + D_i(E^*)]$$

since $\sum_{i \in N} P_i^* = 0$.

(b) Suppose now that the imputation y^* is not in the core. Then, there would exist a coalition S and a partial agreement equilibrium with respect to $S, \tilde{E} = (\tilde{E}_1, \ldots, \tilde{E}_n)$, such that

$$\sum_{i \in S} J_i(\tilde{E}) < \sum_{i \in S} y_i^*. \tag{14}$$

Notice first that $\forall\, i \in N \backslash S, \tilde{E}_i = \bar{E}_i$ in this partial agreement equilibrium because $\bar{E}_i$ is a dominant strategy under linearity of the damage cost functions. Moreover, from the first order conditions that characterize a partial agreement equilibrium one has $\forall\, i \in S, \tilde{E}_i \geq E_i^*$.

Consider now the alternative imputation $\hat{y}$ defined by

$$\hat{y}_i = J_i(E^*) + \hat{P}_i, \quad i = 1, \ldots, n,$$

where the transfers are of the form

$$\hat{P}_i = -[C_i(E_i^*) - C_i(\tilde{E}_i)] + \frac{D_i'}{D_N'} \left[\sum_{i \in N} C_i(E_i^*) - \sum_{i \in N} C_i(\tilde{E}_i) \right]$$

If we can show that

$$\text{(i)} \qquad \sum_{i \in S} \hat{y}_i < \sum_{i \in S} J_i(\tilde{E})$$

and that

$$\text{(ii)} \qquad \sum_{i \in N \backslash S} \hat{y}_i \leq \sum_{i \in N \backslash S} y_i^*,$$

then, given (14), the imputation $\hat{y}$ induces an aggregate cost for all countries which is lower than $w(N)$—an impossibility that proves the theorem.

To show (i), let us write

$$\begin{aligned}
\sum_{i\in S}\hat{y}_i &= \sum_{i\in S} J_i(E^*) + \sum_{i\in S}\hat{P}_i \\
&= \sum_{i\in S} C_i(\tilde{E}_i) + \sum_{i\in S} D_i(E^*) + \frac{\sum_{i\in S} D'_i}{D'_N}\left[\sum_{i\in N} C_i(E^*_i) - \sum_{i\in N} C_i(\tilde{E}_i)\right] \\
&= \sum_{i\in S} C_i(\tilde{E}_i) + \sum_{i\in S} D_i(\tilde{E}) + \frac{\sum_{i\in S} D'_i}{D'_N}\left[D'_N(Q^* - \tilde{Q}) + \left(\sum_{i\in N} C_i(E^*_i) - \sum_{i\in N} C_i(\tilde{E}_i)\right)\right]
\end{aligned}$$

where the last line has been obtained by adding and subtracting $\sum_{i\in S} D_i(\tilde{E})$ to the previous one, and use has been made of the linearity of the functions D_i as well as of the form (6) of the transfer functions. In this expression, a negative value of the term within square brackets can be derived from properties of the optimum E^*, of a partial agreement equilibrium with respect to a coalition and from the strict convexity of the abatement cost functions C_i.

To show (ii), starting from the fact that

$$\begin{aligned}
\hat{y}_i &= C_i(E^*_i) + D_i(E^*) + \hat{P}_i \\
&= C_i(\tilde{E}_i) + D_i(E^*) + \frac{D'_i}{D'_N}\left[\sum_{j\in N} C_i(E^*_i) - \sum_{i\in N} C_i(\tilde{E}_i)\right]
\end{aligned}$$

and

$$\begin{aligned}
y^*_i &= C_i(E^*_i) + D_i(E^*) + P^*_i \\
&= C_i(\bar{E}_i) + D_i(E^*) + \frac{D'_i}{D'_N}\left[\sum_{j\in N} C_i(E^*_i) - \sum_{i\in N} C_i(\bar{E}_i)\right]
\end{aligned}$$

it is sufficient to show that

$$\begin{aligned}
& C_i(\tilde{E}_i) + \frac{D'_i}{D'_N}\left[\sum_{j\in N} C_i(E^*_i) - \sum_{i\in N} C_i(\tilde{E}_i)\right] \\
\leq \quad & C_i(\bar{E}_i) + \frac{D'_i}{D'_N}\left[\sum_{i\in N} C_i(E^*_i) - \sum_{i\in N} C_i(\bar{E}_i)\right], \forall\, i \in N\backslash S.
\end{aligned}$$

But this derives from the characterization of a partial agreement equilibrium with respect to a coalition with linear functions D_i, namely $\tilde{E}_i = \bar{E}_i$, a dominant strategy $\forall\ i \in N\backslash S$, and $\sum_{i\in N} C_i(\tilde{E}_i) \geq \sum_{i\in N} C_i(\bar{E}_i)$ according to Proposition 2.

Q.E.D.
∎

Finally the remark made at the end of Section 4 allows us to further state:

Corollary. The imputation y^* also belongs to the core of the game $[N, w^\alpha]$.

6 Conclusion: Economic Interpretation of the Proposed "γ-core" Solution.

6.1 The Cost Sharing Formula

As announced in the introduction, the essence of the paper is formula (13), which specifies a (net) monetary transfer for each country. Its "core" virtue lies in the fact that, given the international optimum E^*, and the cost (both of abatement and of damage) $J_i(E^*)$ that each country has to bear to implement it, the transfers yield to each country or group of countries an effective net cost lower than the one they would bear under any other arrangement, including under strictly autarkic Nash equilibrium.

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 of formula (13)), and a payment by country i of a proportion D_i'/D_N' of the total of these differences across all countries (second squared bracket of formula (13)). Of course, full knowledge of the functions D_i and C_i is required for computing these amounts. We already had proposed elsewhere (CHANDER and TULKENS 1992) that an international agency be set up to handle these computations and payments. Recall that as they are specified, the transfers break even.

Notice that if $D_i' = 0$ for some country (because it would not be concerned as a recipient—or would allege not to be—of the kind of transfrontier pollution under discussion), while its abatement cost would be positive, that country only receives the first component of the transfer formula, for the abatement it does; but it pays nothing to the others; this leaves it at an effective cost level equal to the one at the Nash equilibrium. In general, though, according to the second component of the formula, the contribution made by a country to the other ones is a fraction of the aggregate abatement costs (above the Nash equilibrium) equal to its relative marginal damage cost to the sum of all countries' marginal damage costs, D_i'/D_N'. In other words, each country's contribution is determined by the relative intensity of its preferences for the public good component of the problem.

6.2 On the Cooperative and Voluntary Nature of Agreements Supported by Core Theoretic Considerations

A core imputation is to be interpreted as a proposal made to all players for sharing $w(N)$, having the property that no coalition S can improve upon it for the benefit of its members by means of the payoff it can secure them. Because of this property, it is claimed that no coalition S is in a position to object to it.

A core proposal is thus (i) a cooperative one, because it involves the all players set, and (ii) one that should entail voluntary agreement, because it is based on an explicit proof that neither the absence of agreement, nor any other agreement can do better for any of the parties involved.

These seem to be strong virtues for designing the cost sharing arrangements in an international treaty.

6.3 On the Role of International Transfers

The strategic role of monetary transfers appears clearly as soon as one realizes that polluting countries with nonzero abatement cost and weak preferences for removal of ambient pollutants never have an interest in cooperating towards abatement (let alone an optimal one), because the costs are higher to them than the benefits. This is reflected, by the way, in the two following remarks: (i) in our game, the core is empty in general if transfers are not allowed; and (ii) when all players are identical, the (unique) core imputation is the one without transfers. It is thus, in a sense, the diversity of the agents that commands transfers when strategic considerations are at stake.

6.4 On the Linearity Assumption on the Damage Cost Function

While restrictive at a general level, the linearity assumption on preferences (i.e., damage costs) may be seen as a mild one in the specific context of transfrontier pollution. On the one hand, at the empirical level, these functions are indeed extremely difficult to estimate. Mainly for that reason, MÄLER 1989-93 and several of his followers have been satisfied with that assumption, all the more that, as he shows, it can be given a useful role in a Nash equilibrium situation.

Another argument is that the optimum may lie far away from the situation prevailing at the time the negotiations begin, thereby increasing uncertainties. Techniques of economic computation have therefore been devised to move *towards* the optimum in successive steps (a recent example is given in GERMAIN, TOINT and TULKENS 1994). For the "local" information required to apply these techniques, linear functions (with parameters possibly varying over time) are surely not inappropriate.

References

BARRETT, S. 1992, "International environmental agreements as games," chapter 1 (pp. 11-36) in Pethig, R. (ed.), *Conflicts and cooperation in managing environmental resources*, Springer-Verlag, Berlin.

CARRARO, C. and SINISCALCO, D. 1993, "Strategies for the international protection of the environment," *Journal of Public Economics* 52, 309-328.

CHANDER, P. 1993, "Dynamic procedures and incentives in public good economies," *Econometrica* 1341-1354.

CHANDER, P. and TULKENS, H. 1992, "Theoretical foundations of negotiations and cost sharing in transfrontier pollution problems," *European Economic Review*, 36 (2/3), 288-299.

CHANDER, P. and TULKENS, H. 1994, "The core of an economy with multilateral environmental externalities," California Institute of Technology Social Science Working Paper 886, August 1994.

d'ASPREMONT, C. and GERARD-VARET, L. 1992, "Incentive theory applied to negotiations on environmental issues: national strategies and international cooperation," mimeo, August.

FOLEY, D. 1970, "Lindahl solution and the core of an economy with public goods," *Econometrica* 38, 66-72.

FRIEDMAN, J. 1990, Game theory with applications to economics, second edition, Oxford University Press.

GERMAIN, M. TOINT, Ph. AND TULKENS, H. 1994, "Calcul économique itératif pour les négociations internationales sur les pluies acides entre la Finlande, la Russie l'Etonie," miméo, Center for Operations Research and Econometrics (CORE), Université Catholique de Louvain.

HOEL, M. 1992, "Emission taxes in a dynamic international game of CO_2 emissions," chapter 2 (pp. 39-68) in Pethig, R. (ed.), *Conflicts and cooperation in managing environmental resources*, Springer-Verlag, Berlin.

LAFFONT, J. J. 1977, *Effects externes et théorie économique*, Monographies du Séminaire d'Econométrie, Editions du Centre national de la Recherche Scientifique (CNRS), Paris.

MÄLER, K. G. 1989-93, "The acid rain game," Chapter 12 (pp. 231-252) in H. Folmer and E. Van Ierland (eds.), *Valuation methods and policy making in environmental economics*, Elsevier, Amsterdam; "The Acid Rain Game II," *Beijer Discussion Paper Series n°32*, Beijer International Institute of Ecological Economics, The Royal Swedish Academy of Sciences, Stockholm.

SCARF, H. 1971, "On the existence of a cooperative solution for a general class of N-person games," *Journal of Economic Theory* 3, 169-181.

SHAPLEY, L. and SHUBIK, M. 1969, "On the core of an economic system with externalities," *American Economic Review* LIX, 678-684.

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*, Gordon and Breach Science Publishers, New York.