International Tax and Public Finance, 2: 279-293 (1995)

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

PARKASH CHANDER
California Institute of Technology, Pasadena, California, and Indian Statistical Institute (ISI), New Delhi, India, e-mail: chander@isid.ernet.in

HENRY TULKENS
Center for Operations Research and Econometrics (CORE), Université Catholique de Louvain, 34 Voie du Roman Pays, B-1348 Louvain-La-Neuve, Belgium, e-mail: corsec@core-ucl.sc.be

Abstract

For a simple economic model of transfrontier pollution, widely used in theoretical studies of international treaties bearing on joint abatement, we offer in this paper a scheme for sharing national abatement costs through international financial transfers that is inspired by a classical solution concept from the theory of cooperative games—namely, the core of a game. The scheme has the following properties: total damage and abatement costs in all countries are minimized (optimality property), and no coalition or subset of countries can achieve lower total costs for its members by taking another course of action in terms of emissions or transfers, under some reasonable assumption about the reactions of those not in the coalition (core property). In the concluding section economic interpretations of the scheme are proposed, including its connection with the free-riding problem.

Key words: cost of abatement, damage function, free riding, partial agreement Nash equilibrium, γ and α cores, optimal emission policy, international transfers

1. Introduction

This 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 (in Section 5) for allocating abatement costs between the countries involved for which the justification is of core-theoretic nature.

We examine this concept because of the need to ensure more than mere optimality in the outcomes of international negotiations leading to a treaty. On the one hand, the optimum sought for should be a voluntary one because of the nonexistence of a supranational authority endowed with sufficient coercive power to impose any emissions policy on countries, even if optimal. On the other hand, the optimum should be robust against the temptation of free riding by some of the countries (or groups of countries) involved, given the public good (actually public bad) nature of transfrontier pollution.

We claim in this paper that a core-theoretically based argument enjoys the two properties just stated and also offer with formula (13) an explicit policy that implements them.

We develop our arguments in the framework of the simplest model traditionally used for the economic analysis of transfrontier pollution agreements (as, for example, in Mäler, 1989–1993; Hoel, 1992; Barrett, 1992; Carraro and Siniscalco, 1993; d'Aspremont and Gerard-Varet, 1992). Our claims in this paper proceed from results that are presented with full technical details in a companion theoretical paper (Chander and Tulkens, 1994), where use is made of a more general model expressed in terms of an Arrow-Debreu economy with public goods, initially formulated in Tulkens (1979) and also used in Chander and Tulkens (1992). The present paper shows that the Chander and Tulkens (1994) results are readily applicable to the more common simple environmental model just evoked.

2. Transfrontier pollution: The economic model and its associated games

2.1. *The basic economic model*

Currently, the model most commonly used for the economic analysis of international agreements on transfrontier pollution is based on the following components:

1. $N = \{i \mid i = 1, \ldots, n\}$, the set of countries concerned by the analysis, the countries being n in number, each indexed by i;
2. For each country i:
   - Quantities $E_i \geq 0$ of some pollutant *emitted* by the economic agents of country i, per unit of time. E_i is a scalar as the analysis bears on one pollutant only. An extension to several pollutants would require E_i to be a vector.
   - Quantities $Q_i \geq 0$ of *ambient* pollutant present in country i's environment, per unit of time. As for the emissions, Q_i is a scalar; with several pollutants it would be a vector.
   - A *transfer function*

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

     where $E = (E_1, \ldots, E_n)$, that describes the physical, chemical, or biological processes whereby the amounts E of pollutants 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.
   - 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 (that is, abating) emissions is costly.
   - 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 ($D_i' \geq 0$).
   - The *total of abatement and damage costs*

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

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.

3. Finally, the description of this international economy with pollution—henceforth summarily designated by the pair $[N, (C_i, D_i, F_i)_{e \in N}]$—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 national costs. E^* is thus the solution of the optimization problem

$$\underset{\{E_1,\ldots,E_n\}}{Min} \quad J(E), \tag{3}$$

where

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

2.2. *The associated games*

As stated in the introduction, 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. 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 (1) by characterizing as equilibria of a *noncooperative game* associated with the above economic model, national emission policies that satisfy only the objectives of each country; and (2) 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,

- A *noncooperative game*, defined by its players set $N = \{i \mid i = 1, \ldots, n\}$, 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 $T_i = \{E_i \mid E_i \geq 0\}$ (with possibly an upper bound E_i^0 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.
- A *cooperative game* (in characteristic function form and with transferable utility), defined by its players set $N = \{i \mid i = 1, \ldots, n\}$ and the function $w(S)$ that associates with every subset S of N a number called the worth of S (or the payoff to coalition 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 players set with the set of countries, and by defining the characteristic function as[1]

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

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 on the strategies $(E_j)_{j \notin S}$ adopted by the countries that 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 and characterization of optimality*

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 $E_i^0 > 0$ such that*

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

Assumption 2. *For all i, the function $C_i(E_i)$ is strictly convex (i.e., $C_i'' > 0$) over the range $]0, E_i^0[$.*

Assumption 3. *For all 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 $Q_i = Q_j$ for all $i, j \in N$, 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.[2] In view of Assumption 3, we shall often write, with some notational inconsistency, $D_i(Q_i)$ as $D_i(E)$.

Assumption 4. *For all 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 for all i, the total cost function $J_i(E)$ is convex. One can then prove, as in Chander and Tulkens (1994, Sec. 3), the following:

Proposition 1. *Under the Assumptions 1 through 4, the optimal joint emissions vector E^* is unique and in the range $]0, E_i^0[$ for all i.*

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

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

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:

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*

$$\forall i,\ \bar{E}_i \text{ minimizes } J_i(E), \text{ where for each } j \in N,\ j \neq i,\ E_j = \bar{E}_j.$$

Existence of this equilibrium follows from standard theorems (see, for example, Friedman, 1990). It is characterized by the first-order conditions

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

As these differ from the conditions in (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 conditon (9) imply

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

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 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 Chander and Tulkens (1994).

3.2. *Strong and coalition-proof equilibria*

Mäler (1989–1993) has considered stronger concepts of voluntary behavior in the framework of noncooperative games—namely, the strong Nash equilibrium and the coalition-proof Nash equilibrium.[3] However, the former is shown by Mäler not to exist in general in the case of the game associated with our economic model, and the latter is not Pareto-efficient, if it exists 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 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*

1. *$(\tilde{E}_i)_{i \in S}$ minimizes $\Sigma_{i \in S} J_i(E)$, where for every $j \in N, j \notin S$, $E_j = \tilde{E}_j$ as defined in* 2 *below, and*
2. *$\forall j \in N \setminus S$, $\tilde{E}_j$ minimizes $J_j(E)$, where for every $i \in S$, $E_i = \tilde{E}_i$ as defined in* 1 *above.*

In Chander and Tulkens (1994, sec. 3.3), we prove the following:

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

1. *There exists a partial-agreement equilibrium with respect to S;*
2. *The vector of individual emission levels at such an equilibrium is unique;*
3. *The individual emissions of the players outside the coalition are not lower than those at a Nash equilibrium;*
4. *The total emissions level is not higher than at a Nash equilibrium.*

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

$$\sum_{j \in S} D_j'(\tilde{E}) + C_i'(\tilde{E}_i) = 0, \qquad i \in S$$

$$D_j'(\tilde{E}) + C_j'(\tilde{E}_j) = 0, \qquad j \in N \setminus S. \tag{11}$$

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 $\Sigma_{i \in N} y_i = w(N)$. 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

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

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 with Proposition 1, the minimum aggregate cost obtained with $w(N)$ is determined by a unique joint strategy vector $E^* = (E_1^*, \ldots, E_n^*)$, yielding $J_i(E^*)$ for each i and of course $\Sigma_{i \in N} J_i(E^*) = w(N)$. 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 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

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

and the condition

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

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

4.3. Alternative characteristic functions

It was observed at the end of Section 2.2 that for arguments $S \neq N$ of the characteristic function (4) 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 member of S and of the other players. To this effect, we shall consider the following two alternatives:

1. The cooperative game $[N, w^{\alpha}]$, defined by the characteristic function of the form

$$w^{\alpha}(S) = \underset{(E_i)_{i \in S}}{Min} \sum_{i \in S} J_i(E) \text{ where, if } S \neq N,\ E_j = E_j^0\ \forall j \in N \backslash S.$$

(Recall that E_j^0 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 that is worst for it—that is, pollute up to E_j^0 in our model.

This form has been often used in economic models with beneficial externalities or with public goods production (see, for example, Foley, 1970; Scarf, 1971; and recently Chander, 1993) where it is a natural one because the worst strategy of nonmember 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^0 for the individual emissions. The worst then becomes infinite amounts, which is hardly credible.

While Mäler concludes by dismissing the tool of the characteristic function, and as a consequence the core conept that is built on it,[4] we choose to propose instead to consider the following alternative:

2. The cooperative game $[N, w^{\gamma}]$, defined by the characteristic function of the form

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

(Recall that $\tilde{E}_j$ was defined in part 2 of Definition 2—definition of a partial-agreement equilibrium with respect to a coalition.)

The function w^{γ} is to be called the *partial agreement equilibrium 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[5] to the actions of S.

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

5. An imputation in the γ-core (and in the α-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 (1969) well known garbage game.[6]

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 outselves here to the first case and refer the reader to Chander and Tulkens (1994) for the second one.

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

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

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

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

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

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

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

since

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

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

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

Notice first that $\forall i \in N \setminus S$, $\tilde{E}_i = \bar{E}_i$ 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

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

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

$$\sum_{i \in S} \hat{y}_i < \sum_{i \in S} J_i(\tilde{E}) \tag{15}$$

and that

$$\sum_{i \in N \setminus S} \hat{y}_i \leq \sum_{i \in N \setminus S} y_i^*, \tag{16}$$

then, given (14), the imputation $\hat{y}$ induces an aggregate cost for all countries that is lower than $w(N)$, the solution of (3)—an impossibility that proves the theorem.

To show (15), let us write

$$\sum_{i \in S} \hat{y}_i = \sum_{i \in S} J_i(E^*) + \sum_{i \in S} \hat{P}_i$$

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

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

where the last line has been obtained by adding and subtracting $\Sigma_{i \in S} D_i(\tilde{E})$ 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 w.r.t. a coalition and from the strict convexity of the abatement cost functions C_i.

To show (16) starting from the fact that

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

and

$$\begin{aligned} y_i^* &= C_i(E_i^*) + D_i(E^*) + P_i^* \\ &= C_i(\bar{E}_i) + D_i(E^*) + \frac{D_i'}{D_N'}\left[\sum_{i \in N} C_i(E_i^*) - \sum_{i \in N} C_i(\bar{E}_i)\right], \end{aligned}$$

it is sufficient to show that

$$C_i(\tilde{E}_i) + \frac{D_i'}{D_N'}\left[\sum_{i \in N} C_i(E_i^*) - \sum_{i \in N} C_i(\tilde{E}_i)\right] \le C_i(\bar{E}_i) + \frac{D_i'}{D_N'}\left[\sum_{i \in N} C_i(E_i^*) - \sum_{i \in N} C_i(\bar{E}_i)\right], \ \forall i \in N \backslash S.$$

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

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 $J_i(E^*)$ (both of abatement and of damage) that each country has to bear to implement it, the transfers yield to each country an effective net cost lower than the one they would bear under any partial agreement, 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* each country i of a proportion D_i'/D_N' of the total of these differences across all countries (second squared bracket of formula (13)). We already have 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 of the transfrontier pollution under discussion—or would allege not to be) while its abatement cost C_i' is positive, that country then only receives the first component of the transfer, based on the abatement it does, and it pays nothing to the others; this leaves the country 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 others is a fraction of the aggregate abatement costs (of moving from the Nash equilibrium to the optimum), the fraction being determined by D_i'/D_N'—that is, its relative marginal damage cost to the sum of all countries' marginal damage costs. 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. *Information requirements*

The core-theoretic property of the proposed solution was established in the framework of games of complete information. To put it in practice, the computation of the optimum E^* as well as of the associated transfers P_i^* stated in (13) requires full knowledge of the abatement- and damage-cost functions in all countries.

As this information has to be provided by the countries themselves, there may be strong incentives for them to give false information. While theoretical analyses of this incentives problem are numerous in the literature,[7] extending the present analysis in that direction is beyond the scope of the present paper. Let us observe, however, that in practice the problem can be eased to some extent by international inspections and audits. Mäler (1989–1993) and Kaitala, Mäler, and Tulkens (1995), for instance, use abatement-cost functions estimated on the basis of some plots produced by the Acid Rain project at IIASA. They also show how damage-cost functions can be calibrated from abatement-cost functions.

6.3. *On cooperation and robustness against free riding*

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 on 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. In the present context of an international externality, objecting against a proposed agreement may be seen as attempting to free ride against it.

Our claim of robustness against free riding for the γ-core solution is in the following sense: given the solution proposed to N, the all players set, if some coalition S envisages to free ride by seeking an arrangement of its own, the breaking up of the players not in S into singletons acting rationally is sufficient to make this free riding less attractive to the members of S than the proposed solution. The threat we assume on the part of the non-free riders is thus the source of the deterrence to free ride; it is also what induces full cooperation. The essence of our contribution is in identifying that rational behavior on the part of singletons is sufficient for coooperation in that sense.

6.4. 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 toward abatement (let alone an optimal one) because the costs are higher to them than the benefits. This strategic aspect is also reflected in the two following remarks: (1) in our game, the core is empty in general if transfers are not allowed, and (2) 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.

Apart from strategic considerations, the payments involved by these transfers (the second bracketed term in formula (13)) are more in the spirit of the "victim pays" principle than of the "polluter pays" principle. This only reflects the fact that the ethical values that inspire the latter are here in opposition to the self-interest considerations that are called on to ensure voluntariness in cooperation and deter free riding.

Finally, we show in Chander and Tulkens (1994) that the solution y^* can also be interpreted as an equilibrium concept, analogous to the one of ratio equilibrium and therefore of Lindahl equilibrium in public-goods economies. Elaborating on this requires, however, the Arrow-Debreu setting used in that paper.

6.5. On the linearity assumption on the damage-cost function

While restrictive at a general level, the linearity assumption on preferences (that is, 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 (1989–1993) 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 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 *toward* 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.

Acknowledgments

This paper was prepared for the Fiftieth Congress on Public Finance, Environment and Natural Resources of the International Institute of Public Finance, held at Cambridge, Massachusetts. Thanks are due to Karl Göran Mäler for numerous fruitful discussions and his hospitality at the Beijer International Institute for Ecological Economics, Royal Swedish Academy of Sciences, Stockholm in May 1993, to Tito Cordella for a very careful reading, as well as to Jack Mintz and Carlo Carraro, discussants at the Congress. The first author is also grateful to the 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) project on Environmental Policy, International Agreements and International Trade administered by Professor Alistair Ulph through CEPR, London.

Notes

1. 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 Chander and Tulkens (1994) allows for payoffs to be directly defined on the utilities of the economic agents, more in line with usual practice.
2. In Mäler's (1989–1993) 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 $\Sigma_{i \in N} a_{ji} = 1$, the externalities are directional and do not have the public good property.
3. Both concepts have been discussed in detail in Bernheim, Peleg, and Whinston (1987).
4. Another argument made is that with no bounds on the behaviors 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.
5. 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 $S \setminus \{i\}$ 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.
6. 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.
7. Kwerel (1977), and Dasgupta, Hammond, and Maskin (1980) propose schemes in which truthful revelation is a dominant strategy but that do not balance the budget.

References

Barrett, S. (1992). "International Environmental Agreements as Games." In Pethig, R. (ed.), *Conflicts and Cooperation in Managing Environmental Resources* (pp. 11–36). Berlin: Springer-Verlag.

Bernheim, D., B. Peleg, and M.D. Whinston. (1986). "Coalition-Proof Nash Equilibria. I: Concepts." *Journal of Economic Theory* 42(1), 1–12.

Carraro, C., and D. Siniscalco. (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 H. Tulkens. (1992). "Theoretical Foundations of Negotiations and Cost Sharing in Transfrontier Pollution Problems." *European Review* 36(2/3), 288-299.

Chander, P., and H. Tulkens. (1994). "The Core of an Economy with Multilateral Environmental Externalities." Mimeo, forthcoming *CORE Discussion Paper*, April.

Dasgupta, P., P. Hammond, and E. Maskin. (1980). "On Imperfect Information and Optimal Pollution control." *Review of Economic Studies* 47, 857-860.

d'Aspremont, C., and L. Gerard-Varet. (1992). "Incentive Theory Applied to Negotiations on Environmental Issues: National Strategies and International Cooperation." Mimeo, August.

Dreze, J.H., and D. de la Vallee Poussin. (1971). "A Tâtonnement Process for Public Goods." *Review of Economic Studies* 38, 133-150.

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* (2nd ed.). Oxford: Oxford University Press.

Germain, M., Ph. Toint, and H. Tulkens. (1994). "Calcul économique itératif pour les négociations internationales sur les pluies acides entre la Finlande, la Russie, l'Etonie." Mimeo, Center for Operations Research and Econometrics (CORE), Catholic University of Louvain.

Hoel, M. (1992). "Emission Taxes in a Dynamic International Game of CO_2 Emissions." In R. Pethig (ed.), *Conflicts and Cooperation in Managing Environmental Resources* (pp. 39-68). Berlin: Springer-Verlag.

Kaitala, V., K.G. Mäler, and H. Tulkens. (1995). "The Acid Rain Game as a Resource Allocation Process with Application to Negotiations Between Finland, Russia and Estonia." *Scandinavian Journal of Economics* 97(2).

Kwerel, E. (1977). "To Tell the Truth: Imperfect Information and Optimal Pollution Control." *Review of Economic Studies* 44, 595-601

Laffont, J.J. (1977). *Effets externes et théorie économique*. Paris: Monographies du Séminaire d'Econométrie, Editions du Centre National de la Recherche Scientifique (CNRS).

Mäler, K.G. (1989-1993). "The Acid Rain Game," chapter 12. In H. Folmer and E. Van Ierland (eds.), *Valuation Methods and Policy Making in Environmental Economics* (pp. 231-252. Amsterdam: Elsevier. "The Acid Rain Game II." *Beijer Discussion Paper Series* 32, Beijer International Institute of Ecological Economics, 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 M. Shubik. (1969). "On the Core of an Economic System with Externalities." *American Economic Review* 59, 678-684.

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