

Working Paper 93-25
Economics Series 06
October 1993

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9875

ENVIRONMENTAL CONSCIOUSNESS AND MORAL HAZARD
IN INTERNATIONAL AGREEMENTS TO PROTECT THE
ENVIRONMENT

Emmanuel Petrakis and Anastasios Xepapadeas*

Abstract

A group of countries that can potentially commit to cooperation to protect the environment are identified as environmentally-conscious countries. Conditions are examined under which they can provide self-financing side payments, to a second group of less environmentally-conscious countries, so that the two groups form a global or partial stable coalition that agrees to emit at the first-best global welfare optimum. A mechanism is also developed which must be incorporated into the agreement between the two groups, in order to induce all countries to emit at the desired level, even when global pollution has nonpoint source pollution characteristics.

Key Words

Environmental Economics, Moral Hazard in Teams.

*Petrakis, Departamento de Economía, Universidad Carlos III de Madrid; Xepapadeas, Departments of Economics, University of Crete, Greece. We thank the audiences of the 1993 Econometric Society European Meeting, Uppsala, Sweden, the 8th Annual Congress of the European Economic Association, Helsinki, Finland, and the seminar of the Center of Planning and Economic Research, Athens, Greece for their comments. Finally, special thanks to Sofia Katsulaki for bringing to our attention the article of the Civil Code on which is based the construction of our mechanism.

ENVIRONMENTAL CONSCIOUSNESS AND MORAL HAZARD IN INTERNATIONAL AGREEMENTS TO PROTECT THE ENVIRONMENT

1. Introduction

The international dimension of environmental policy arises from the fact that pollutants generated as an inevitable byproduct of production in any country can cross national borders, carried by physical media like water or air. As a result, it is the total amount of accumulated pollutants that affects a country and not just the country's own emissions. These international environmental problems vary in nature from truly global problems, such as the depletion of the ozone layer or the global warming associated with the greenhouse effect, to more localized problems, such as the pollution of the Mediterranean or the acid rain problem in Europe. In practice, attempts to form a policy capable of dealing with international problems have focused on international agreements (e.g., the Helsinki Protocol or the Montreal Protocol). In the recent literature in environmental economics, the focus has been on the emergence of international cooperation and underlying incentives (Carraro and Siniscalco 1992a).

The basic question in the case of cooperation is whether sovereign countries can voluntarily (since there is no authority to force their cooperation) reach an agreement to protect the international commons. When this type of cooperation is examined, there is a major difficulty since each country has an incentive to defect from the agreement when everybody else cooperates, since by doing so a country can reduce the cost of abating pollution and enjoy the benefits from the overall reduced pollution brought about by the cooperation of the rest of the countries. Furthermore, as has been shown by Carraro and Siniscalco (1992b), a small number of cooperating countries can not enlarge the coalition

with self-financed utility transfers to countries that have no incentive to join it, unless the cooperating countries are committed to signing the agreement and then providing incentive for the enlargement of the cooperating group. This paper follows that argument.

A group of countries that can potentially commit to cooperation are identified as the environmentally-conscious countries (ENCC). By assuming that the ENCC are committed to an agreement to protect the environment, conditions are examined under which these countries can provide side payments, which are self-financing, to a second group of less environmentally-conscious countries (LENCC). Under the proposed side payments scheme, all countries form a stable coalition that agrees to adjust their emissions so that a first-best global welfare optimum is achieved. A side payment scheme capable of securing a stable partial coalition of the ENCC with a subset of the LENCC is also examined.

Having determined the system of side payments, a second problem is examined related to difficulties in enforcing the agreement, due to moral hazard. In particular, the case is considered in which the level of global pollution can be measured at a relatively low cost, whereas the cost of measuring each individual country's contribution to global pollution is prohibitively costly for an external observer. The country itself, on the other hand, can measure its contribution to global pollution at a relatively low cost. The inability to measure each country's emissions creates problems in the enforcement of the agreement, since countries have an incentive to cheat by emitting more than the agreement dictates. This is, therefore, a type of nonpoint source pollution case with an international dimension. Although a nonpoint source problem arising within a country can be handled by the imposition of collective penalty schemes (Segerson 1988, Xepapadeas 1991), the same formula does not apply here since there is no supranational authority to impose collective penalty schemes on sovereign countries. Thus in a second stage, a mechanism is developed which must be

incorporated into the agreement between the ENCC and the LENCC and which will induce the latter countries to emit at the desired level even when global pollution has nonpoint source characteristics.

The rest of the paper is organized as follows. Section 2 analyzes conditions for the achievement of a total or partial stable coalition; section 3 presents the mechanism that secures the enforcement of the agreement when monitoring of individual country's emissions is not possible; and section 4 provides some concluding remarks.

2. First-Best Optimum and Stable Coalitions

We consider a group of $i=1, \dots, N$ countries. Production activities in each country generate pollution. Pollution crosses boundaries and its total concentration in the environment affects the whole group. We assume that the benefit of each country is a strictly concave function of its emissions, e_i , defined as:

$$B_i(e_i), B_i(0)=0, B_i' > 0, B_i'' < 0$$

where e_i belongs to some compact and convex subset of R_+ . The ambient concentration of the pollution defined as $X = \sum_i e_i$ causes damages to all N countries. We assume that the damage function in each country takes the simple linear form $m_i X$, where $m_i \geq 0$ measures marginal damages in country i . Following Hoel (1992), we assume that:

$$m_1 \geq m_2 \geq \dots \geq m_N, M = \sum_i m_i > 0$$

Therefore, countries with a low index, or high m , are considered to suffer the most damages from environmental pollution. If $m=0$ for some country, then this country does not suffer any damages from global pollution. Under these assumptions, the benefit function for

country i is defined as a function of all other countries' emission levels:

$$b_i(e_i, e_{-i}) = B_i(e_i) - m_i X \quad (1)$$

where $e_{-i} \equiv (e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_N)$. If there is no international cooperation, each country's emissions are determined as the Nash equilibrium of the non-cooperative game where each country chooses emissions simultaneously to maximize (1). Thus the Nash equilibrium emission level for each country is defined as:

$$\hat{e}_i: B_i'(\hat{e}_i) - m_i = 0, \quad i=1, \dots, N \quad (2)$$

The non-cooperative emission levels can be compared to the emission levels corresponding to the first-best optimum which is obtained by maximizing global welfare, defined as the sum of individual countries' benefit functions:

$$V(e_1, \dots, e_N) = \sum_i b_i(e_i, e_{-i}) \quad (3)$$

A solution to the maximization of (3) can be regarded as the case where a supranational authority, if it existed, would set the emission levels for each country at the first-best optimum. This first-best emission level for each country is defined as:

$$\bar{e}_i: B_i'(\bar{e}_i) - M = 0, \quad i=1, \dots, N \quad (4)$$

Assuming that at least two m s are non-zero, then from (2) and (4), we have that:

$$\bar{e}_i < \hat{e}_i \quad \forall i, \quad \bar{X} < \hat{X}$$

This is a well-established result in the analysis of the international aspects of pollution (e.g., van der Ploeg and de Zeeuw 1992). Not all countries, however, are better off at the first-best optimum as compared to the Nash equilibrium (Hoel 1992). Some countries for which:

$$m_i < \frac{B_i(\bar{e}_i) - B_i(\hat{e}_i)}{\bar{X} - \hat{X}} \quad (5)$$

will be worse off at the first-best optimum in comparison to the Nash equilibrium. These countries will be characterized by low m in the case where the benefit functions of all countries are fairly similar. On the other hand, countries for which:

$$m_i > \frac{B_i(\bar{e}_i) - B_i(\hat{e}_i)}{\bar{X} - \hat{X}} \quad (6)$$

will be better off at the first-best optimum. We define countries for which (6) holds as environmentally conscious, and countries for which (5) holds as less environmentally conscious. We index the ENCC by $j \in J$ and the LENCC by $k \in K$. The sets J and K are assumed to be not empty in order to make the problem interesting. They partition the set $I = \{1, \dots, N\}$, that is, $J \cup K = I$, $J \cap K = \phi$.

It is clear that in the presence of ENCC and LENCC, there will be no voluntary cooperation by all N countries for the achievement of the first-best. Assume that the ENCC can form a coalition under which each country will emit at the first-best level. This coalition will be stable if it is both internally and externally stable in the sense of the cartel stability defined in d'Aspremont *et al.* (1983). Internal stability means that there is no incentive to defect from the coalition, or:

$$B_j(\hat{e}_j) - B_j(\bar{e}_j) < m_j(\hat{e}_j - \bar{e}_j), \quad \forall j \in J. \quad (7)$$

It should be noted that although environmental consciousness, as defined in (6), is necessary for internal stability, it is not, however, sufficient. Thus the ENCC can not always form an internally-stable coalition. External stability, on the other hand, means that there is no incen-

tive for any LENCC to enter the coalition and emit at the first-best level, or:

$$B_k(\hat{e}_k) - B_k(\bar{e}_k) > m_k(\hat{e}_k - \bar{e}_k), \quad \forall k \in K. \quad (8)$$

Conditions are now examined under which this coalition can achieve the first-best optimum. Assume that the ENCC try to expand the coalition by self-financing side payments to the LENCC. In order to attract a country from the fringe of the LENCC, say country k , the total benefits of the ENCC after attracting country k must exceed the loss of this country, or:

$$\sum_{j \in J} m_j(\hat{e}_k - \bar{e}_k) > B_k(\hat{e}_k) - B_k(\bar{e}_k) - m_k(\hat{e}_k - \bar{e}_k) \quad (9)$$

When (9) holds and country k receives a transfer greater than or equal to its loss, country k has an incentive to join the coalition. If this happens, however, and the transfer is paid, some members of the ENCC coalition might want to defect. This can be shown from (7). If the amount that j has to contribute to induce country k to enter the coalition is sufficiently high so that the inequality in (7) is reversed, then country j has an incentive to defect. This underlying inability to enlarge the coalition through self-financing transfers so that new members emit at the first-best is similar to results obtained by Carraro and Siniscalco (1992b). It is interesting to note that the more environmentally conscious a country is, the stronger the incentive will be to defect from the coalition.

A stable larger coalition could, however, be formed if some subset of the group of countries were to commit to cooperation (Carraro and Siniscalco 1992b). Conditions are examined, therefore, under which a commitment to cooperation by the ENCC, to emit at the first-best, could induce cooperation by all LENCC so that the whole group is a stable coalition. These conditions are stated in the following proposition.

Proposition 1:

(i) If all J countries commit to cooperation, and (ii) if

$$\sum_j [B_j(\bar{e}_j) - B_j(\hat{e}_j) - m_j(\bar{X} - \hat{X})] \geq \sum_k [B_k(\hat{e}_k) - B_k(\bar{e}_k) - m_k(\hat{e}_k - \bar{e}_k)], \quad (10)$$

then there exists a system of self-financed side payments $t_k, k \in K$ such that all LENCC enter the coalition and emit at the first-best level of emission. The coalition of all N countries is stable.

Proof

The k th country from the LENCC will be no worse off under the first-best if it receives a side payment:

$$\hat{t}_k = B_k(\hat{e}_k) - B_k(\bar{e}_k) - m_k(\hat{X} - \bar{X})$$

which is equal to its losses for moving to the first-best. This transfer might not, however, be enough to ensure stability of the coalition, since country k will have an incentive to defect once all other countries have entered the coalition. The gains from defection are defined as:

$$d_k = [B_k(\hat{e}_k) - m_k(\hat{e}_k + \sum_{i \neq k} \bar{e}_i)] - [B_k(\bar{e}_k) - m_k \sum_i \bar{e}_i] \quad \text{or} \quad (11)$$

$$d_k = B_k(\hat{e}_k) - B_k(\bar{e}_k) - m_k(\hat{e}_k - \bar{e}_k)$$

If $t_k \geq d_k$, then there will be no incentive for country k to defect. Thus the coalition is internally stable since by (i) the J countries are committed to cooperation and also by (ii) the side payments are self-financed. Q.E.D.

The above proposition implies that if there is a large number of ENCC so that the benefits from moving to the first-best are sufficient, then the ENCC can offer sufficient inducement for the LENCC to form a stable coalition wherein all countries will emit at the

first-best. A stable coalition can also be achieved if the marginal pollution damages of the LENCC are sufficiently small. To show that, write:

$$\hat{e}_k = \hat{X} - \sum_{\substack{s \in I \\ s \neq k}} \hat{e}_s, \quad \bar{e}_k = \bar{X} - \sum_{\substack{s \in I \\ s \neq k}} \bar{e}_s$$

Substituting \hat{e}_k and \bar{e}_k into (10) and rearranging terms, we obtain:

$$\sum_{i=1}^N [B_i(\bar{e}_i) - B_i(\hat{e}_i) - m_i(\bar{X} - \hat{X})] \geq \sum_k m_k \left(\sum_{\substack{s \in I \\ s \neq k}} \hat{e}_s - \sum_{\substack{s \in I \\ s \neq k}} \bar{e}_s \right) \quad (12)$$

The right-hand side of (12) is always positive, since it represents the difference in total benefits between the first-best and the noncooperative Nash equilibrium. Therefore (12) or equivalently (10) will be satisfied if m_k are sufficiently small.

The requirement that all countries enter the coalition might be too strong. However, the approach used to derive conditions that ensure the stability of the coalition of all countries can be used to examine partial expansion. By a partial expansion, it is meant that the ENCC and some of the LENCC can form a stable coalition with all countries emitting at the first-best optimum, while the rest of the LENCC emit at the noncooperative Nash equilibrium. Assume that the set K of LENCC is partitioned into sets H and Q , $h \in H$, $q \in Q$, $Q \neq \phi$. Assume that countries in set H emit at the first-best optimum while countries in set Q emit at the noncooperative Nash level, and define:

$$X' = \sum_j \bar{e}_j + \sum_h \bar{e}_h + \sum_q \hat{e}_q, \text{ then clearly } \bar{X} < X' < \hat{X}.$$

Assume finally that for all ENCC in set J , condition (6) is satisfied when \bar{X} is replaced by X' . If (6) is not satisfied for some ENCC when \bar{X} is replaced by X' , the set of ENCC should be redefined accordingly. Then the condition for a partial expansion of the coalition of the

ENCC can be stated as follows.

Proposition 2:

(i) If all ENCC commit to cooperation, and (ii) if

$$\sum_j [B_j(\bar{e}_j) - B_j(\hat{e}_j) - m_j(X' - \hat{X})] \geq \sum_h [B_h(\hat{e}_h) - B_h(\bar{e}_h) - m_h(\hat{e}_h - \bar{e}_h)] \quad (13)$$

then there exists a system of self-financed side payments t_h , $h \in H$, such that the LENCC that belong to set H enter the coalition and emit at the first-best. The coalition of the countries that belong to the set $J \cup H$ is stable.

Proof

The gains from defection for any country $h \in H$ are defined as:

$$d_h = \left[B_h(\hat{e}_h) - m_h \left(\hat{e}_h + \sum_q \hat{e}_q + \sum_j \bar{e}_j + \sum_{\substack{s \neq h \\ s \in H}} \bar{e}_s \right) \right] - \left[B_h(\bar{e}_h) - m_h \left(\sum_j \bar{e}_j + \sum_h \bar{e}_h + \sum_q \hat{e}_q \right) \right] \text{ or} \\ d_h = B_h(\hat{e}_h) - B_h(\bar{e}_h) - m_h(\hat{e}_h - \bar{e}_h) \quad (14)$$

If $t_h \geq d_h$ there will be no incentive for country h to defect. Thus the partially expanded coalition is stable. Q.E.D.

By comparing (13) with (10), it can be noted that although the gains of the ENCC from the partial expansion of the coalition, and thus the total amount available for side payments, is less than the case where a total coalition is achieved, a successful partial expansion might be possible if there are LENCC for which the marginal pollution damage is sufficiently small. This can be seen by rewriting (13) in order to obtain an inequality similar to (12).

The achievement of the first-best requires, however, that the emissions of each individual country can be observed. If this is not possible, either because it is too costly or because it is technically impossible, then the LENCC will have an incentive to cheat. An

additional mechanism is required to ensure the enforcement of the agreement so that the first-best or a partial enlargement of the initial coalition of ENCC is achieved. This is the problem addressed in the next section.

3. Enforcing the Agreement under Moral Hazard

If monitoring of the emission level of an individual country is prohibitively costly, then the ENCC, which in the sequel are supposed to act as a single agent because they are committed to cooperation, can not observe the contribution to the global pollution by each of the LENCC. On the other hand, an LENCC k , if asked, is able to verify its emission level by incurring a fixed cost $f_k > 0$, $k \in K$. This asymmetry between the costs of monitoring and verification can easily be justified. An external observer sent by the ENCC to measure the emission level of an LENCC will most probably spend a lot of money to collect the information required for the report, while the same information could be collected at a relatively low cost by the country itself. It is further assumed that the cost of measuring global pollution is low, and to simplify things, it is set at zero.

The question addressed in this section is whether the global optimum (or the optimal level of pollution in case of a partial agreement) can still be achieved in the presence of *moral hazard*; that is, when each LENCC has an incentive to first sign the agreement and then overemit. If a supranational authority existed, it could impose collective penalty schemes to eliminate any incentives of an LENCC to cheat, but this is not possible in an international setup. To deal with moral hazard in the international commons, a modification of a procedure which is an integral part of the Civil Code of many countries¹, and has been applied repeatedly in the courts, is proposed. This procedure mandates that if a Debtor A

¹See, for example, section IV, chapter III, book IV of the Spanish Civil Code.

owes a given amount of money to a group of agents, without knowing how much he owes to each one of them, he has the right to arbitrarily choose one agent (say, his creditor 1) and transfer the whole amount of debt to him. Creditor 1 then becomes a debtor towards the rest of the group. Creditor 1, in turn, may arbitrarily choose one of the other creditors and transfer the whole amount of debt to him, minus the part he claims A owes to him. This procedure continues up to the point where only one creditor is left; he then receives the residual. If the last creditor is not satisfied with the residual, he may appeal to the courts asking for more money from his predecessor. If the last creditor proves his claim to be valid, he may in turn appeal to the courts, accusing the previous creditor in the chain of a false claim, and so on.

Consider the following game. The players are the LENCC, each acting in its own interest, and the group of ENCC acting as a single agent. The game is played in *three stages*. The first is the *negotiation* stage. In this stage all parties involved sign an agreement which specifies (i) each individual country's level of emission x_i^* , $i=1, \dots, N$, (ii) a vector of lump-sum transfers from the ENCC to the LENCC, t_k , $k \in K$, with $T = \sum_k t_k$, to be paid if the measured global pollution, X_m , equals the agreed $X^* = \sum x_i^*$; and (iii) a total amount of transfers, $T_c = T - \min_k t_k$, for every other $X_m \neq X^*$ to be distributed among the LENCC through the *Mechanism (M*)*, which is described below. The second is the *actual pollution* stage. Each country decides independently its emission level x_i , $i = 1, \dots, N$. Given that by assumption $x_j = x_j^*$ for all $j \in J$, if $X_m \leq X^*$ each LENCC receives t_k and the game ends. If, however, $X_m > X^*$, that is, if one or more of the LENCC have overemitted, we proceed to the last stage of the game, the *implementation of the Mechanism M** stage.

Note that whenever the measured pollution level is higher than the agreed, by no matter how much, the *same* total amount of transfers will be distributed among the LENCC.

This amount is equal to the sum of the transfers minus its smallest component, if all parties respect the agreement. This implies that there is always enough money for transfers to those LENCC which stick to the agreement. Put otherwise, given that an LENCC which cheats has no right to receive any side-payment, the ENCC "owe" at most T_c to the group of LENCC if $X_m > X^*$. The *Mechanism M** then works as follows.

The ENCC randomly select one of the LENCC (w.l.o.g., say LENCC 1) and make it a "promise" of transfer T_c . LENCC 1 in turn chooses at random one of the other LENCC (say 2) and makes it a "promise" of $T_c - s_1$, thus claiming s_1 for itself. If $s_1 \neq t_1$, then LENCC 1 will receive a zero transfer at the end of the procedure, while the ENCC "debt" is automatically reduced by t_1 . An LENCC which claimed $s_1 \neq t_1$ must also announce its level of emissions. LENCC 2 then receives a "promise" of $T_c - t_1$, either because LENCC 1 claimed t_1 or because t_1 has subtracted from ENCC "debt". The procedure is repeated by LENCC 2 replacing LENCC 1 and so on. In this way, the last LENCC in the chain (say ℓ) always receives a "promise" less than t_ℓ , since $\sum_{k \in K} t_k = T > T_c$, $k \in K$. If the sum of the announced overemissions of those LENCC which claimed $s_k \neq t_k$ equals $X_m - X^*$ (measured minus agreed level of pollution), the game ends with the ENCC satisfying all the outstanding claims as well as making a transfer t_ℓ to LENCC ℓ .

Otherwise, LENCC ℓ has two options: to initiate the *procedure of verification* or not. If not, it receives a zero transfer in the end and must also announce its emission level. The procedure of verification works backwards. Starting from the last in the chain, LENCC ℓ asks *sequentially* for verification of the emission level from its predecessors which still have outstanding claims. That is, LENCC ℓ will request verification first from LENCC $\ell-h$ if LENCC $\ell-1, \dots, \ell-h+1$ have claimed $s_k \neq t_k$, and so on. An LENCC asked for verification is obliged to provide it. If its emission level differs from the agreed one, it receives a transfer

of zero in the end. The procedure of verification stops whenever the sum of overemissions of both LENCC which verified and those which initially claimed $s_k \neq t_k$ equals $X_m - X^*$. Then all still outstanding claims are paid and LENCC l receives t_l . Finally, if LENCC l does not start the procedure of verification, its first predecessor with an outstanding claim, LENCC $l-h$ replaces it, and again faces the above two options; and so on.

Note that an LENCC has nothing to gain from a false announcement of its emission level, so we shall assume that such announcements are truthful. Further, an LENCC which overemitted but still starts the procedure of verification, will be discovered eventually (and will receive no transfer) because the sum of overemissions will turn out to be less than $X_m - X^*$ even after all LENCC have announced or verified their emission levels. Thus an LENCC which overemitted has no gains from the verification procedure.

To show that the global optimum can be achieved, suppose that $t_k = d_k$ with d_k defined in (11). An LENCC receives a transfer equal to its gains from defection from the first-best whenever it emits at the first-best level, or overemits but is not caught. Then if condition (10) is satisfied and ENCC are committed to cooperation, all countries will be better off at the global optimum, provided that no LENCC has an incentive to cheat by overemitting. Hence, if the above mechanism detects cheating, all countries have an incentive to sign the proposed scheme in the negotiation stage. We now prove that the above scheme does, in fact, detect cheating.

Proposition 3:

Given that $x_j = e_j$ for $j \in J$, $x_k = e_k$ $k \in K$ is a Perfect Equilibrium for the Subgame starting from stage 2. Equivalently, each LENCC will independently choose its emission level according to the first-best.

Proof: First, an LENCC k which emits at the agreed level e_k will always make a claim of

t_k if it can, or else it will start the procedure of verification if necessary, despite whatever actions the other LENCC have taken so far, and despite whatever LENCC k 's beliefs are about the future actions of the other LENCC. Suppose now that $x_k = \bar{e}_k$ for all $k \in K$, $k \neq \ell$. Does LENCC ℓ have an incentive to choose $x_\ell \neq \bar{e}_\ell$? For any $x_\ell \leq \bar{e}_\ell$, $X_m \leq X^*$ and LENCC ℓ will receive t_ℓ ; further, $b_\ell'(x_\ell, \dots) < 0$ for all $x_\ell \leq \bar{e}_\ell$. Thus LENCC ℓ has no incentive to underemit. Does it have an incentive to overemit? If $x_\ell > \bar{e}_\ell$, then $X_m > X^*$ and mechanism M^* will be implemented. Given the fact that LENCC ℓ can infer that it is the only one which overemitted, the best it can do is not to start the procedure of verification if it is last in the chain, and to claim any $s_\ell \neq t_\ell$ in all other positions. If it comes last in the chain and it starts the procedure of verification, it will be caught eventually because all others will be able to verify their emission level. Then it receives the same transfer of zero as if it had not started it. If it asks for t_ℓ in any other position, it knows that the last LENCC in the chain will start the verification procedure, so LENCC ℓ will be asked to verify its emission level. Not only does it receive a zero transfer in this case, but it also incurs a cost f_ℓ . By claiming $s_\ell \neq t_\ell$ or not starting the verification procedure, however, LENCC ℓ will end up receiving a zero transfer. Thus, an argument similar to proposition 1 shows that LENCC ℓ has no incentive to deviate from \bar{e}_ℓ . This shows that $x_k = \bar{e}_k$, $k \in K$ is a Nash Equilibrium of the subgame starting from stage 2.

To show that this a Perfect Equilibrium, it is necessary to specify LENCC ℓ strategy in the case where it infers that it is not the only country which overemitted. The best strategy that LENCC ℓ can follow is never to start the verification procedure, and whenever asked to claim $s_\ell \neq t_\ell$, regardless of whatever actions other LENCC have taken so far and whatever ℓ 's beliefs about future actions of the remaining LENCC. Note first that all LENCC, which either claimed $s_k \neq t_k$ or did not start the verification procedure, have also

announced their emission levels. Thus if LENCC ℓ starts the verification procedure (either being the last in the chain or having received the initiative from the next LENCC to it), it will be caught eventually because even after all the rest of the LENCC have verified their actual emission levels, the measured global pollution will still exceed the sum of individually announced and verified levels. Hence it cannot improve upon a zero transfer. If it claims t_ℓ , there will always be some country which will start the verification procedure. If the initiative is given to LENCC ℓ , we showed that it will pass it to the LENCC preceding it, thus receiving a zero transfer. If some other LENCC starts the verification procedure, LENCC ℓ will be asked to verify its emission level, thus spending f_ℓ and moreover receiving no transfer. Again it cannot improve upon a zero transfer. This completes the proof that $x_k = \bar{e}_k$, $k \in K$ is a Perfect Equilibrium. Q.E.D.

Note that in equilibrium no country will ever be asked to verify its emission level, thus no cost f_k will ever be incurred. Notice further that there might be cases out of equilibrium where the system of transfers is not self-financed. However, it could be assumed that the ENCC are committed not only to cooperation, but also to incurring some additional costs if something goes wrong. Proposition 3 allows us to extend proposition 1 in the presence of moral hazard.

Proposition 4:

If (i) all J ENCC commit to cooperation, and (ii) condition (10) holds, there exists a system of self-financed side payments t_k , $k \in K$ and a mechanism (M^) such that each LENCC has an incentive to emit at its first-best emission level even if individual monitoring is prohibitively costly.*

Similar results can be obtained if the required side payments are not sufficient to obtain the cooperation of all countries involved, but instead are enough for a partial

augmentation of the initial coalition of the ENCC.

4. Summary

When each country is affected not just by its own emissions, but by the total amount of global pollution, then international environmental policy should be developed. In this paper, conditions under which sovereign countries can voluntarily achieve an agreement to protect the international commons are examined. A problem which arises in the case of international commons is that each country has an incentive to defect from the agreement when all other countries cooperate, since by doing so a country can reduce the cost of abating pollution and enjoy the benefits from the overall reduced pollution brought about by the cooperation of the rest of the countries.

It is shown that if there exists a group of environmentally-conscious countries - identified as countries gaining from moving from the noncooperative equilibrium to the first-best - which are committed to protecting the environment, these countries can induce all or a subset of the less environmentally-conscious countries to also emit at a global first-best level by a system of self-financed side payments if either of the following conditions holds: (i) the number of ENCC is large enough, and/or their benefits from achieving the global first-best are sufficiently high, or (ii) the marginal disutility of each of the LENCC is sufficiently small. The implicit assumption here is that ENCC act as a single agent, which therefore has incentive to stick to its first-best emission level, and pay the total amount of transfers.

It is further shown that observability of the emission level of each LENCC is not crucial to our results. Even if individual monitoring is prohibitively costly, the global first-best can be achieved under the same conditions as previously. Given that global pollution can

be measured costlessly, and that an LENCC is able to verify its emission level if ever asked to do so, by incurring a cost, a mechanism is developed which guarantees that each LENCC has incentive to report its emission level truthfully. The possibility of errors in measurement of the global or individual levels of emissions has not been examined in this paper. It remains for further research to answer whether a global first-best can be achieved if there exist errors in measuring pollution and/or uncertainty on the part of certain countries in controlling their pollution level.

REFERENCES

Carraro, C. and Siniscalco, D., "The international dimension of environmental policy," *European Economic Review*, **36**, 1992a, 379-387.

Carraro, C. and Siniscalco, D., "The international protection of the environment: Voluntary agreements among sovereign countries," in *The Protection of Trans-National Commons* (P. Dasgupta, K.G. Maler and A. Vercelli, Eds.) Clarendon, Oxford, 1992b.

d'Aspremont, C., Jacquemin, A., Gabszewicz, J., and Weymark, J., "On the stability of collusive price leadership," *Canadian Journal of Economics*, **XVI**(1), 1983, 17-25.

Hoel, M., "International environment conventions: The case of uniform reductions of emissions," *Environmental and Resource Economics*, **2**, 1992, 141-159.

Segerson, Kathleen, "Uncertainty and incentives in nonpoint pollution control," *Journal of Environmental Economics and Management*, **15**, 1988, 87-98.

van der Ploeg, F. and de Zeeuw, A., "International aspects of pollution control," *Environmental and Resource Economics*, **2**, 1992, 117-139.

Xepapadeas, A., "Environmental policy under imperfect information: Incentives and moral hazard," *Journal of Environmental Economics and Management*, **20**, 1991, 113-126.