IKERLANAK

AN APPROACH TO THE STABILITY OF INTERNATIONAL ENVIRONMENTAL AGREEMENTS: THE ABSORBING SETS SOLUTION

by

Norma Olaizola

2003

Working Paper Series: IL. 10/03

Departamento de Fundamentos del Análisis Económico I

Ekonomi Analisiaren Oinarriak I Saila

University of the Basque Country

An Approach to the Stability of International Environmental Agreements: The Absorbing Sets Solution.

Norma Olaizola[*][†]

December 2003

Abstract

We study international environmental negotiations when agreements between countries can not be binding. A problem with this kind of negotiations is that countries have incentives for free-riding from such agreements. We develope a notion of equilibrium based on the assumption that countries can create and dissolve agreements in their seeking of a larger welfare. This approach leads to a larger degree of cooperation compared to models based on the internal-external stability approach.

JEL Classification: Q50, C70, C72.

Key words: International Environmental Agreements, Absorbing Sets Solution.

[*] Dpto. de Fundamentos del Análisis Económico I. Facultad de Ciencias Económicas y Empresariales. University of the Basque Country. Avda. L. Agirre 83, 48015, Bilbao, Spain. Fax: 34 946013891, Phone: 34 946017075. e-mail: etpolorn@bs.ehu.es

[†] The author thanks E. Inarra, J. Kuipers, J. Arin, J.C. Bárcena, M. Escapa and A. Saracho for helpful comments. Financial support from Basque Country University project: 9/UPV 00035.321-13699/2001 and Ministerio de Educación y Ciencia project: BEC2000-0875 is gratefully acknowledged.

1 Introduction

Unsustainable economic activity is putting increasing stress on air, water, energy and other resources: Greenhouse gases threaten to change the climate, chemicals reduce the ozone layer and other pollutants can cause acid rain or travel long distances to cause damage on land and water. These emissions usually affect to the global environment and consequently generate a problem of transfrontier externalities. Governments should put into force international environmental agreements (which we call IEAs) calling for reductions in these harmful substances. But in the absence of an international institution that has the authority to enforce environmental policy, it is difficult to create binding agreements among sovereign countries.

However, as the outcomes of the Montreal Protocol on the deplection of the ozone layer or the Rio and Kyoto Protocols on climate changes show, international environmental agreements do in fact occur. Therefore, although full cooperation is difficult to attain or very unstable, we could expect partial cooperation among countries. The main question in this paper is whether or not it is possible to obtain some degree of cooperation. Moreover we aim to determine which type of agreements among groups of countries are stable and which ones not when confronted to pollution reduction. We propose an approach and a solution concept for analyzing this type of problems which gives rise to full cooperation.[1]

As we have mentioned, there is no an international institution able to force binding agreements. Then, a fundamental assumption will be that the agreements reached by countries must be self-enforcing, that is not based on the countries' commitment to cooperation. But a self-enforcing agreement involving a large number of countries is difficult to attain. In fact, countries have a free-rider incentive: They would be better off by not cooperating and enjoying the pollution abatement efforts of the cooperating countries while not incurring any (or low) pollution abatement costs. This free-riding effect induces a high instability in the agreements, since, once a coalition has been formed, members of the coalition obtain a lower profit than outsiders and thus have an incentive to break the agreement. Therefore, even if cooperation among all countries is the efficient outcome, by no means constitutes an equilibrium.

There exist in the literature different approaches to this type of problems. Barrett [1994] considers a model with identical countries and analyzes whether it is possible that countries reach a self-enforcing agreement. He uses a stability concept borrowed from the industrial organization literature on cartels (d'Aspremont et al. [1983]).[2] This stability concept predicts stable coalitions which involve a small number of countries. While the definition of stability due to d'Aspremont et al. [1983] excludes the possibility of coalitional deviations, Thoron [1998] develops a model of coalition formation in oligopolies which in-

[1] This approach was previously introduced in Inarra, Kuipers and Olaizola [2002].

[2] A stable coalition is one where, for all countries in the coalition, there is no incentive to defect from the coalition (internal stability) and, for all countries outside the coalition, there is no incentive to join the coalition (external stability).

corporates the possibility of coalitional deviations. As in Barrett [1994], only the formation of one coalition at a time is allowed. On the contrary Ray and Vohra [2001] propose a model which allows the formation of multiple coalitions of countries at the same time. They develop a notion of equilibrium: For a coalition to form, all the countries in that coalition must view this coalition as their best choice and each coalition evaluates its payo¤ based in a rational prediction of what the remaining countries will do. They characterize the unique equilibrium coalition structure for this model and also show that the number of coalitions that form in equilibrium is small.

Now let us brie‡y explain our approach. As in Barrett [1994], we allow for the formation of only one coalition of countries. In this work we determine endogenously the number of signatories and the terms of the agreement, that is the magnitude of the pollution reduction of signatories and also of non-signatories. Given an IEA, some countries may be satis…ed but others may try to force the transition to another IEA. In this work we describe how transitions among IEAs take place and take up the question of how countries react if a country or group of countries deviate. We allow countries to move successively in discrete steps unless they converge upon some IEAs which are stable. We describe this process by means of an abstract game. An abstract game consists on a set of elements and a binary relation de…ned on that set. In our case the set of elements represent the set of all IEAs, and the binary relation de…ned on that set represents the transitions between IEAs. Several concepts have been de…ned for abstract games, among them we choose the absorbing sets solution for solving the model. An absorbing set in our context is a set of IEAs satisfying these three conditions: First, whenever the negotiation process leads to an IEA that belongs to an absorbing set, it may shift to another IEA that belongs to the same absorbing set in a …nite number of steps. Therefore, once an IEA in an absorbing set is reached there is a positive probability of transition among all the IEAs in this set. Second, whenever the bargaining process leads to an IEA that belongs to an absorbing set, from this time on any IEA that does not belong to this absorbing set is impossible to reach (even if it belongs to another absorbing set). And third the bargaining process will with certainty lead in a …nite number of steps to an IEA in one of the absorbing sets. Let us see now how the Kyoto protocol operates for justifying why we view the absorbing sets solution as an appropriate solution concept for studying international environmental negotiation models.

For the Kyoto protocol to became a legally binding agreement, it must be rati…ed by not less than 55 countries which must be responsible for a total of at least 55% of the 1990 emissions reported by the group of developed countries and emerging economies. Until this state is reached the protocol is non-binding, therefore countries may either ratify the protocol or withdraw from it. For example the United States signed the protocol but does not intend to ratify it. Russia signed the protocol as well and it is said that it planned to ratify it but recently decided to withdraw. As countries' decisions are highly linked, a reconsideration on the part of a country could provoke further reconsiderations of other countries. Therefore the negotiation process can easily lead to a situation with no a unique stable IEA but a group of 'stable' IEAs that appear and disappear as countries negotiate in their finding of better agreements. Following this procedure we suggest a process of transition from an IEA to another which derives the negotiation to some IEAs while another IEAs are ruled out as stable.

The main result of the work characterizes the equilibrium IEAs. This approach leads to a larger degree of cooperation compared to models based on the internal-external stability approach. In fact, we obtain a unique set of stable IEAs, and this set contains the full cooperative agreement.

The remainder of the paper is structured as follows: The model is presented in Section 2, in which the process through which countries form agreements is described. The stability notion and the results are presented in section 3. Section 4concludes the paper.

2 The Model[3]

Let be a world of $i = 1, \ldots, n$ countries, each of which emits a pollutant that damages a shared environment. We regard a country as an economic agent which has to choose the pollution abatement level that maximizes the total welfare of its economy.

Let x_i denote the part of the total reduction of pollution that can be ascribed to country i, hence the total reduction of pollution is $X = \sum_{i=1}^{n} x_i$. And let $W_i(X)$ be the welfare of country i, which depends on the total reduction pollution level:

$$W_i(X) = B_i(X) - C_i(x_i).$$

$B_i(X)$ function denotes the benefit that country i obtains from global pollution reduction, this benefit depends on the total amount of pollution reduction. In this work we consider constant marginal benefit function:[4]

$$B_i(X) = X.$$

The cost for country i is assumed to be a quadratic function of i's contribution to the pollution reduction. We suppose that all countries have the same pollution reduction technology:

$$C_i(x_i) = \frac{1}{2}x_i^2.$$

The net profit (welfare) for country i is thus:

$$W_i(X) = X - \frac{1}{2}x_i^2.$$

[3] This model is standard in the literature about international environmental agreements formation. See for example Barrett [1994] and Ray and Vohra [2001].

[4] We suppose marginal benefit equal to 1 for simplicity in the exposition. However, the results of the work do not change considering any other value for the marginal benefit.

2.1 Coalitional Cooperation

We consider an International Environmental Agreement (IEA) as a coalition S formed by those countries which decide to cooperate to reduce their own emissions of pollutants, being s the cardinality of S. If this agreement takes place, s countries cooperate while $n - s$ countries remain as free-riders. Assume that countries in coalition S jointly decide their contribution to the problem of pollution reduction taking into account the behavior of the non-signatory countries, while non-signatories maximize their individual welfare. Either signatories and non-signatories choose simultaneously their contributions (Cournot-Nash assumption). Transfers among countries are not allowed. For our model, the aggregate coalitional welfare of coalition S is:

$$\sum_{i \in S} \left(X - \frac{1}{2} x_i^2 \right).$$

The problem of countries in S is to choose the per member level of pollution abatement x_i for all $i \in S$ which maximizes the coalitional welfare taking the abatement levels of the non-signatories as given. That is coalition S solves:

$$\max_{\{x_i\}_{i \in S}} \left(s \sum_{i=1} x_i - \frac{1}{2} \sum_{i \in S} x_i^2 \right).$$

The first order conditions provide:

$$s - x_i = 0 \qquad \text{for all } i \in S.$$

Therefore, all countries in coalition S maximize coalitional profits by setting $x_i = s$.

We can consider the non-cooperating countries as coalitions which give rise to IEAs formed by a unique country. Then, the maximization problem a non-cooperating faces is a particular case of the problem solved by coalition S. Therefore, a non-cooperating country maximizes its profit by setting $x_i = 1$.[5]

The total pollution abatement is thus:

$$X = s^2 - s + n.$$

Now we have the optimal pollution abatement level for all countries, and substituting equilibrium values into their objective functions we obtain the associated equilibrium welfare values. Denote as $W_i(S)$ the final payoff of a signatory country and as $W_i(\neg S)$ the final payoff of a non-signatory when the IEA given by coalition S takes place. Then:

$$W_i(S) = \frac{1}{2} s^2 - s + n \text{ and } W_i(\neg S) = s^2 - s + n - \frac{1}{2}.$$

[5] We are considering constant marginal benefit of pollution abatement equal to 1. In general we would obtain $x_i = b$, being b the marginal benefit.

It is straightforward to see that given any IEA it is better to be a non-signatory country than to sign the agreement. Note also that the welfare of either non-signatories and signatories is larger as more countries add to the IEA that is in force. This last implies that the best agreement for a signatory country is the coalition formed by all countries, N. Straightforward computations show also that a country which withdraws from an IEA (given that the rest of countries do not move) will be better o¤ if and only if the initial agreement includes strictly more than three countries.

Of course, the non cooperative situation where each country i maximizes its own pro…t (no IEA is signed) and the total cooperative solution where all the countries cooperate in the maximization of the global welfare (all countries sign the IEA) are special cases of the coalitional cooperation.

The non-cooperative solution:

In this case $s = 1$. Then, under the non-cooperative outcome, $x_i = 1$ for all $i = 1; :::; n;$ and the global pollution abatement is $X_{:c} = n:$

The cooperative solution:

In this case $s = n$. Then, under the cooperative outcome $x_i = n$ for all $i = 1; :::; n;$ and the global pollution abatement is $X_c = n^2:$

Note that the coalition formed by all countries produces the largest global amount of pollution reduction while the non-cooperative situation produces the lowest global amount of pollution reduction. Note also that the di¤erence, the gains of cooperation, grows as n grows: However it is easy to see that, whenever $n > 3;$ cooperation is unstable. In fact countries have a free-rider incentive: They would be better o¤ by not cooperating and enjoying the pollution abatement e¤orts of the cooperating countries. Taking into account this free-rider incentive, in what follows we study which IEAs are likely to occur.

2.2 Transitions among IEAs

Assume that countries will sign an IEA if they …nd it pro…table. Now suppose that a group of countries form an IEA. Some countries may be satis…ed with the agreement while others may not. Since we suppose that the agreements are not binding, the non satis…ed countries will try to …nd something better for them. We assume that the decision of signing or not an IEA is motivated by the payo¤s that countries can obtain, in our model these payo¤s are provided by the welfare function W_i. We consider that to form an IEA an agreement among all countries involved is needed, while countries may freely leave an IEA. Therefore, no country can be forced to sign an IEA and signatories to an IEA can always withdraw from the agreement.

In what follows we analyze the existence of stable agreements.

2.2.1 Pro…table Transitions

Let S and S^0 be two coalitions of countries and let $M(S; S^0)$ be a minimum set of countries necessary to force the transition from S to S^0. It is easy to see that this minimum set may be not unique. In general we have:

i) If $S' \subset S$ there are two minimum sets of countries necessary to force the transition from S to S': $M(S;S') = \{i \in S \cap S'\}$ and $M'(S;S') = \{i \in S \setminus S'\}$.

ii) If $S' \not\subset S$ there is one minimum set of countries necessary to force the transition from S to S': $M(S;S') = \{i \in S'\}$.

However, not all countries in such a minimum set are necessarily better off if the transition from S to S' takes place. This fact leads to the following definition.

Definition 1 Let S and S' be two coalitions. And let $M(S;S')$ be a minimum set necessary to force the transition from S to S'. We say that $M(S;S')$ is a transition makers set of the transition from S to S', and denote $T(S;S')$, if no country in $M(S;S')$ loses with the transition from S to S' and at least one country in $M(S;S')$ strictly gains with this transition.

We denote as $\mathcal{T}(S;S')$ the collection of all the transition makers sets of the transition from S to S'. The following definition states that the transition from S to S' is profitable whenever there is at least one transition makers set for this transition.

Definition 2 Let S and S' be two coalitions. We say that the transition from S to S' is profitable, and denote $S \to S'$, if $\mathcal{T}(S;S') \neq \emptyset$.

If the transition from S to S' is not profitable we denote it as $S \nrightarrow S'$.

2.2.2 Strongly profitable transitions

We have defined what a profitable transition is, but not all the profitable transitions seem equally reasonable. The following definition formalizes this idea.

Definition 3 We say that the transition from S to S' is strongly profitable, and denote $S \overset{s}{\to} S'$, if
i) the transition from S to S' is profitable and
ii) there is no another profitable transition from S, say to S'', such that for any $T(S;S') \in \mathcal{T}(S;S')$ there is one $T(S;S'') \in \mathcal{T}(S;S'')$ such that every $i \in T(S;S') \cap T(S;S'')$ gets at least as much if S'' forms than if S' forms and at least one $i \in T(S;S'')$ is strictly better off if S'' forms than if S' forms.

If the transition from S to T is not strongly profitable we denote it as $S \overset{s}{\nrightarrow} S'$.
Some useful relationships between profitable and strongly profitable transitions can be stated:

Lemma 1[6] Let S be any coalition such that $S \to N$. Then $S \overset{s}{\to} N$.

[6] All proofs are postponed to the appendix.

Lemma 2 Let S be any coalition and S' be the coalition obtained by deleting one country from S. If $S \rightarrow S'$ then either $S \stackrel{S}{\rightarrow} S'$ or $S \stackrel{S}{\rightarrow} N$.

Based on the strongly profitable transitions we next define a dominance relation between IEAs.

Definition 4 We say that coalition S' dominates coalition S if and only if the transition from S to S' is strongly profitable.

In this way, the information provided by the welfare function W_i suggests a process of transition from an IEA to another, which we identify with the strongly profitable transitions. This transitional process derives the negotiation to some IEAs while another IEAs will be ruled out as stable. In what follows we determine the number of signatories in the stable IEAs.

3 Stable IEAs

In this section we analyze the stability of IEAs. We are interested in finding the possible choices that countries would make among the different available IEAs. In doing that, we use the absorbing sets solution for abstract systems[7] which we define in the following.

Definition 5 An Abstract System is a pair $(S; R)$ where S is a set of alternatives and R is a binary relation defined over that set.

Let the pair $(S; R)$ be an abstract system. For $a, b \in S$, aRb means that a dominates b.

A path from a to b in S is a sequence of alternatives $a = a_0, a_1, a_2, \ldots a_m = b \in S$ such that $a_{i-1} R a_i$ for all $i \in \{1, \ldots, m\}$.

Let R^T be the transitive closure of R (i.e. aR^Tb means that there is a path from a to b).

Definition 6 Let $(S; R)$ be an abstract system. A nonempty set $A \subseteq S$ is called an Absorbing Set if:

i) For all $a, b \in A$ $(a \neq b)$: aR^Tb

ii) There is no $b \in S \setminus A$ and $a \in A$ such that bR^Ta.

The absorbing sets solution for an abstract system is the collection of all its absorbing sets.

In our case the set of alternatives S is the set of all possible IEAs and the binary relation R is given by the strongly profitable transitions[8].

[7] The absorbing sets solution is the collection of all the absorbing sets. Each absorbing set coincides with the elementary dynamic solution introduced by Shenoy [1979]. The union of all elementary dynamic solutions is called dynamic solution. This solution was previously defined by Kalai et al. [1976] under the name of the admissible set. Schwartz [1974] introduces also an equivalent solution.

[8] Then aRb is equivalent to $b \stackrel{S}{\rightarrow} a$.

Definition 6 tell us that each of the absorbing sets satisfy two conditions. The first condition says that for any two alternatives in an absorbing set one dominates the other, if not directly then through a path. This implies that whenever the negotiation process leads to an IEA that belongs to an absorbing set, it may shift to any other IEA that belongs to the same absorbing set in a finite number of steps. Therefore, once an IEA in an absorbing set is reached there is a positive probability of transition among all the IEAs in this set. The second condition says that no alternative outside the absorbing set dominates an alternative in the set, even through a path. This implies that whenever the negotiation process leads to an IEA that belongs to an absorbing set, from this time on any IEA that does not belong to this absorbing set is impossible to reach (even if it belongs to another absorbing set). Moreover, the bargaining process will, with certainty, lead in a finite number of steps to an IEA in one of the absorbing sets. Thus, the binary relation defined by the strongly profitable transitions describes the reasoning that countries use in the bargaining process that results in an IEA. Next we determine the stable IEAs for our model.

Define $\bar{s}$ as the minimum number of countries needed to form an IEA such that every non-signatory country receives a profit at least as large as the profit it would obtain if full cooperation forms. Then we can formalize $\bar{s}$:

$$\bar{s} = \min\{s : W_i(\neg S) \geq W_i(N)\}.$$

For our welfare function we have: $\bar{s} = \left\lceil \frac{1}{2} + \frac{1}{2}\left(2n^2 - 4n + 3\right)^{\frac{1}{2}} \right\rceil$. Then we have the following result:

Proposition 1 *The unique absorbing set is $A = \{S \in \mathcal{S}$ such that $s \geq \bar{s} - 1\}$.*

Note that $\bar{s} > 3$ for all $n > 3$. Comparing this result with Proposition 2 in Barrett [1994], which says that the self-enforcing IEA consists of 2 countries when $n = 2$ and 3 countries when $n \geq 3$, we see that our stability notion allows for a larger degree of cooperation than Barrett's stability notion. As $\bar{s}$ is increasing in n, the larger the number of countries involved, the higher the difference in the degree of cooperation between our result and Barret's result. Note also that, according to our stability notion, the full cooperative agreement is always stable.

4 Conclusions

Since for our specific welfare function more cooperation produces larger levels of pollution reduction, a criterion for comparison between the results of different models can be the concentration of cooperation in equilibrium. The concept of internal-external stability leads to rather pessimistic results with respect to international cooperation on pollution control. Barrett [1994] studies different environmental problems characterized by different functional specifications of benefits and costs of pollution abatement. He shows that, depending on the functional speci…cations, a stable international environmental agreement may not exist, or it may not be able to sustain more that two or three signatory countries. Our stability concept, although does not specify a unique stable IEA, leads to a larger degree of stability. Moreover the full cooperative agreement is a stable outcome.

A limitation of our analysis is the use of a speci…c functional form for the environmental welfare function. In the obtaining of our results we use constant marginal bene…ts and linear marginal costs, but we could use di¤erent functional forms for bene…ts and costs. In Barrett [1994] we …nd a casuistry for di¤erent functional forms: Linear marginal abatement bene…ts and costs, constant marginal bene…ts and logarithmic marginal costs, linear marginal bene…ts and constant marginal costs. But he does not get a higher degree of cooperation than in the case of constant marginal bene…ts and linear marginal costs. Only for the case of linear marginal abatement bene…ts and costs the IEA can sustain a large number of signatories, but only when the di¤erence in net bene…ts between the noncooperative and the full cooperative outcomes is very small. Therefore the improvement upon the noncooperative outcome is small. It would be interesting to analyze if the absorbing sets solution predicts also more cooperation than the internal-external stability based solution for these other cases.

Another limitation of the analysis is the assumption of symmetry between countries. In fact environmental problems are often characterized by large asymmetries across countries. For the cases of climate change and deplection of ozone layer, although CO_2 and CFCs emissions a¤ect all the countries, valuation of the environmental damage can be very di¤erent among, for example, developed and developing countries. The risk of being a¤ected by deplection of ozone layer is high for South Chile and Argentina. The e¤ect of climate change can be also very di¤erent, some island and coastal countries as the Maldives or the Nederland may be specially damaged by the rising of see level due to climate warming while other countries as the former Soviet Union or Canada could bene…t. Such di¤erences may have a large in‡uence on the ability to sustain an international environmental agreement. In this context, Hoel [1992] analyzes a problem of negotiating self enforcing environmental agreements among heterogeneous countries. He assumes marginal abatement bene…ts constant but di¤erent across countries and marginal cost function linear and identical for all countries. Using the same stability concept than Barrett [1994] he …nds that the self-enforcing agreement consists of at most three countries.

A full characterization of the absorbing sets solution under the assumption of heterogeneity between countries is di¢cult to obtain but some partial results are available upon request.

5 Appendix

Proof of Lemma 1 The unique transition makers set of $S \to N$ is $T(S;N) = \{i \in N\}$. Then, we have to prove that there is no T such that $S \to T$ and such that there is one $T(S;T) \in T(S;T)$ such that every $i \in T(S;T)$ gets at least as much if T forms than if N forms and at least one $i \in T(S;T)$ is strictly better off if T forms than if N forms. Consider first $S \to T$ with $T \subset S$. There are two minimum sets that could force the transition from S to $T \subset S$: $M(S;T) = \{i \in S \cap T\}$ and $M'(S;T) = \{i \in S \setminus T\}$. As the welfare of a signatory increases as cooperation increases, then $W_i(T) < W_i(N)$ for any T. On the other hand we know that $W_i(\neg T) < W_i(\neg S)$ and since $S \to N$ then $W_i(\neg S) < W_i(N)$. Therefore, $W_i(\neg T) < W_i(N)$ for all $T \subset S$. Consider now $S \to T$ with $T \not\subset S$. Then, the unique minimum set that could force the transition from S to $T \not\subset S$ is $M(S;T) = \{i \in T\}$. As the welfare of a signatory increases as cooperation increases then all countries in $M(S;T)$ are strictly better off if N forms than if T forms. ■

Proof of Lemma 2 As $S' \subset S$, the unique transition makers set for $S \to S'$ is $M(S;S') = \{l\}$, being l the unique country in $S \setminus S'$.

i) First we prove that $(S \to S'$ and $S \overset{S}{\not\to} S') \Rightarrow S \overset{S}{\to} N$: As $S \overset{S}{\not\to} S'$, then there exists a coalition T such that $S \to T$ and such that country l is not worse off if T forms than if S' forms. It is easy to see that $T \not\subset S$: Now we prove that one such a coalition is coalition N: The unique minimum set that could force the transition from S to N is $M(S;N) = \{i \in N\}$. We know that $\max W_i(S) = W_i(N)$. And since $S \subset T$, then $(S \to S'$ and $\overset{S}{\not\to} S') \Rightarrow S \to N$. Finally by Lemma 1 we have that $S \to N \Rightarrow S \overset{S}{\to} N$:

ii) Now we prove that $(S \to S'$ and $S \overset{S}{\not\to} N) \Rightarrow S \overset{S}{\to} S'$: Suppose $S \overset{S}{\not\to} S'$, then there exists a coalition T such that $S \to T$ and such that country l is not worse off if T forms than if S' forms. We will show that this leads to a contradiction. Consider first $T \not\subset S$. The unique minimum set that could force the transition from S to $T \not\subset S$ is $M(S;T) = \{i \in T\}$. By Lemma 1 we know that $S \to N \Rightarrow S \overset{S}{\to} N$: Then we have that $S \not\to N$. As $\max W_i(S) = W_i(N)$, then $S \not\to N \Rightarrow S \not\to T$ for any $T \not\subset S$. Now consider $T \subset S$. The unique minimum set that could force the transition from S to $T \subset S$: $M(S;T) = \{i \in S \setminus T\}$. As that the welfare of a non-signatory increases as cooperation increases, then $W_l(\neg S') > W_l(\neg T)$ for all $T \subset S'$. Therefore there is no T such that $S \to T$ and such that l is not worse off if T forms than if S' forms. Thus, $S \overset{S}{\to} S'$: ■

Proof of Proposition 1 Let us consider the following sequence of coalitions: $S_0; S_1; ...; S_k; ...; S_{n-\overline{s}+1}$; being S_k the coalition that results if we delete k countries from coalition N. As $\overline{s} > 3$ we have that $S_{k-1} \to S_k$ for all $k = 0; ...; n-\overline{s}+1$ with $T(S_{k-1}; S_k) = \{l\}$; being l any signatory in S_{k-1}. Now we show that for all T such that $S_{k-1} \to T$ and for all k, country l is nnot worse off if S_k forms than if T forms and thus, $S_{k-1} \overset{S}{\to} S_k$ for all $k = 0; ...; n - \overline{s} + 1$. As the welfare of a non-signatory is larger as cooperation increases, then $S_{k-1} \overset{S}{\to} S_k \Rightarrow$

$S_{k-2} \stackrel{s}{\to} S_{k-1}$. Therefore we only need to prove $S_{n-\bar{s}} \stackrel{s}{\to} S_{n-\bar{s}+1}$. Lemma 2 says that $S_{n-\bar{s}} \to S_{n-\bar{s}+1} \Rightarrow$ either $S_{n-\bar{s}} \stackrel{s}{\to} S_{n-\bar{s}+1}$ or $S_{n-\bar{s}} \stackrel{s}{\to} N$. By construction of $S_{n-\bar{s}}$ we have that $W_i(S_{n-\bar{s}+1}) > W_i(N)$. Since it is better to be a non-signatory than to sign an agreement, then $W_i(\neg S_{n-\bar{s}}) > W_i(N)$. As the unique minimum set that could force the transition from $S_{n-\bar{s}}$ to N is $M(S_{n-\bar{s}}, N) = \{i \in N\}$; then $S_{n-\bar{s}} \nrightarrow N$ and thus, $S_{n-\bar{s}} \stackrel{s}{\to} S_{n-\bar{s}+1}$. By construction of S_m we have that $W_i(\neg S_{n-\bar{s}}) \leq W_i(N)$ and as it is better to be a non-signatory than to sign an agreement, then $S_{n-\bar{s}} \to N$. Finally by Lemma 1 we have that $S_{n-\bar{s}} \to N \Rightarrow S_{n-\bar{s}+1} \stackrel{s}{\to} N$. On the other hand, $S_m \to N$ for any $m < n - \bar{s}$. Then by Lemma 1 we have $S_m \stackrel{s}{\to} N$ for any $m < n - \bar{s}$. ■

6 References

[1] Barrett S., 1994, Self-enforcing International Environmental Agreements. Oxford Economic Papers: 46, 878-894.

[2] D' Aspremont C.A, A. Jacquemin, J.J. Gabszewicz and J. Weymark, 1983, On the stability of collusive price leadership. Canadian Journal of Economics: 16, 17-25.

[3] Hoel M., 1992, International Environmental Conventions: The case of Uniform Reductions of Emissions. Environmental and Resource Economics: 2, 141-159.

[4] Inarra E., J. Kuipers and N. Olaizola, 2002, Absorbing Sets in Coalitional Systems. Biltoki, 2002.01.

[5] Inarra E., J. Kuipers and N. Olaizola, 2004, Absorbing and Generalized Stable Sets. Social Choice and Welfare: Forthcoming.

[6] Kalai E., E.A. Pazner and D. Schmeidler, 1976, Admissible Outcomes of Social Bargaining Processes as Collective Choice Correspondence. Journal of Economic Theory: 63, 299-325.

[7] Ray D. and R. Vohra, 1997, Equilibrium Binding Agreements. Journal of Economic Theory: 73, 30-78.

[8] Ray D. and R. Vohra, 2001, Coalitional Power and Public Goods. Journal of Political Economy: 109, 1355-1384.

[9] Shenoy L., 1979, On Coalition Formation: A Game Theoretical Approach. International Journal of Game Theory: 8, 133-164.

[10] Schwartz T., 1974, Notes on the Abstract Theory of Collective Choice. School of Urban and Public A¤airs, Canegie-Mellon University.

[11] Thoron S., 1998, Formation of a coalition-proof stable cartel. Canadian Journal of Economics: 31, 63-76.