June 2018

Fondazione Eni Enrico Mattei

Working Paper

022.2018

International Environmental Agreements - The Impact of Heterogeneity among Countries on Stability

Effrosyni Diamantoudi, Eftichios Sartzetakis, Stefania Strantza

Climate and Sustainable Innovation
Series Editor: Massimo Tavoni

International Environmental Agreements - The Impact of Heterogeneity among Countries on Stability

By Effrosyni Diamantoudi, Department of Economics, Concordia University
Eftichios Sartzetakis, Department of Economics, University of Macedonia
Stefania Strantza, Department of Economics, Concordia University

Summary

The present paper examines the stability of self-enforcing International Environmental Agreements (IEAs) among heterogeneous countries in a two-stage emission game. In the first stage each country decides whether or not to join the agreement, while in the second stage the quantity of emissions is chosen simultaneously by all countries. We use quadratic benefit and environmental damage functions and assume k types of countries that differ in their sensitivity to the global pollutant. We find that the introduction of heterogeneity does not yield larger stable coalitions. In particular, we show that, in the case of two types, when stable coalitions exist their size is very small, and, if the asymmetry is strong enough, they include only one type of countries. Moreover, heterogeneity can reduce the scope of cooperation relative to the homogeneous case. We demonstrated that introducing asymmetry into a stable, under symmetry, agreement can disturb stability.

Keywords: Environmental Agreements

JEL Classification: D6, Q5, C7

Address for correspondence:
Effrosyni Diamantoudi
Department of Economics, Concordia University
1455 De Maisonneuve Blvd West
Montreal Quebec H3G 1M8
Canada
E-mail: effrosyni.diamantoudi@concordia.ca

International Environmental Agreements - The Impact of Heterogeneity among Countries on Stability

Effrosyni Diamantoudi* Eftichios Sartzetakis†
Concordia University University of Macedonia

Stefania Strantza‡
Concordia University

April 2018

Abstract

The present paper examines the stability of self-enforcing International Environmental Agreements (IEAs) among heterogeneous countries in a two-stage emission game. In the first stage each country decides whether or not to join the agreement, while in the second stage the quantity of emissions is chosen simultaneously by all countries. We use quadratic benefit and environmental damage functions and assume k types of countries that differ in their sensitivity to the global pollutant. We find that the introduction of heterogeneity does not yield larger stable coalitions. In particular, we show that, in the case of two types, when stable coalitions exist their size is very small, and, if the asymmetry is strong enough, they include only one type of countries. Moreover, heterogeneity can reduce the scope of cooperation relative to the homogeneous case. We demostrated that introducing asymmetry into a stable, under symmetry, agreement can disturb stability.

Keywords: Environmental Agreements, *JEL*: D6, Q5, C7

*Corresponding Author: Effrosyni Diamantoudi, Department of Economics, Concordia University, 1455 De Maisonneuve Blvd West, Montreal Quebec H3G 1M8, Canada (effrosyni.diamantoudi@concordia.ca).

†Department of Economics, University of Macedonia, 156 Egnatia Str., Thessaloniki 54006, Greece.

‡Department of Economics, Concordia University, 1455 De Maisonneuve Blvd West, Montreal Quebec H3G 1M8, Canada.

Figure 1

Acknowledgement

"This paper has received funding from the European Union's Horizon 2020 research and innovation programme Marie Sklodowska-Curie Actions (IF), under grant agreement No 706620"

1 Introduction

The most important environmental problem, that of climate change, has an international dimension and thus, it can only be addressed effectively through international cooperation. However, the absence of a supranational authority that could enforce environmental actions on sovereign states necessitates the development of self enforcing agreements. Such an agreement should provide countries with incentives to join and remain as members. A country will join an agreement if its net benefits as a signatory exceed its net benefits as a non-signatory. D'Aspremont et al. (1983) introduced a concept of coalitional stability whereby no member has an incentive to leave (internal stability) and no non-member has an incentive to join (external stability), assuming that the rest of the agents do not change their membership decision. The notion is essentially a Nash equilibrium where the strategy choice is to join a coalition or not.

The main body of the literature models the formation of IEAs as a two-stage non-cooperative game: in the first stage countries decide whether to join the coalition, while in the second they choose their emission level depending on their membership status. In the second stage, it is assumed that either all countries (signatories or not) choose emissions simultaneously or that the coalition acts as a leader and the non-signatories follow[1]. The subgame perfect Nash equilibrium of the resulting two-stage game is usually derived by applying the notions of the aforementioned internal and external stability conditions.

Although it is clear that all countries benefit from cooperation, each country has strong incentives to free ride on the coalition's efforts. Free-riding incentives increase as the costs of reducing emissions increase. The literature shows that the size of a stable coalition is small, regardless of the total number of countries. Assuming quadratic cost and benefit functions and simultaneous choice of emissions, it has been shown that stable coalitions consist of no more than two countries (De Cara and Rotillon, 2001; Finus and Rundshagen, 2001; and Rubio and Casino 2001; among others). If the coalition is assumed to be a leader, a stable coalition could have more than two members, but still a maximum of four countries (Barrett, 1994; Diamantoudi and Sartzetakis, 2006)[2].

[1] The two approaches yield similar results.

[2] Barrett 1994 suggests that a stable coalition may achieve a high degree of cooperation, including the grand coalition, but only when an accumulation of stock pollutant is assumed and

One of the most restrictive assumptions of the literature so far is the homogeneity of countries' costs and benefits. It is widely accepted that both damages suffered from a global pollutant and benefits derived from emitting the pollutant (related to production and consumption) differ significantly among countries. The present paper addresses this issue by introducing k types of countries that differ in their sensitivity to the global pollutant. We find that the introduction of heterogeneity does not yield larger stable coalitions. In particular, we show that, in the case of two types, the internal stability condition holds only for coalitions with maximum two members from each type of countries. Furthermore, the external stability condition holds only for coalitions consisting of one type of countries, if the asymmetry is strong. Only for very small asymmetry, a mixed coalition consisting of one country from each type is stable. Finally, we demonstrate that coalitions that are stable under asymmetry they become unstable when asymmetry is introduced. Therefore, the assumption of homogeneity is not the determining factor driving the pessimistic result of small stable coalitions.

Despite its apparent importance, only a few papers have addressed the issue of heterogeneity within a theoretical framework, albeit in a limited way. Assuming two types of countries, Barrett (1997) finds no substantial difference in the size of the stable coalition relative to the homogeneous case. On the contrary, McGinty (2007), allowing for transfer payments through a permit system, finds that heterogeneity can increase the coalition size. Chou and Sylla (2008) consider two types of countries (denoted developed and developing) and provide a theoretical framework to explain why it is more likely that some developed countries form a small stable coalition first and then engage in monetary transfers to form the grand coalition. Osmani and Tol (2010) assume also two types of countries but allow the formation of two separate coalitions. They demonstrate that in the case of high environmental damages, forming two coalitions yields higher welfare and better environmental quality relative to a unique coalition. Biancardi and Villani (2010) introduce asymmetry in environmental awareness and find that the coalitions' stability depends on the level of the asymmetry and that the grand coalition can be obtained only by transfers. Fuentes-Albero and Rubio (2010) assume that countries differ either in abatement costs or environmental damages (which are assumed to be linear on

therefore per period abatement can exceed per period emissions. In contrast, Diamantoudi and Sartzetakis (2006) demonstrate that when no stock pollutant is present and emissions must be positive (interior solution) the stable coalition cannot have more than four members.

emissions) and find that heterogeneity has no important effect without transfers, but if transfers are allowed the level of cooperation increases with the degree of heterogeneity. Finally, Pavlova and Zeeuw (2013) assuming differences in both emission-related benefits and environmental damages (which are assumed to be linear on emissions), find that large stable coalitions are possible without transfers if the asymmetries are sufficiently large, however, the gains of cooperation are very low. As the above review indicates, results of the theoretical literature are mixed. Most of the literature based on simulations finds that under some circumstances, heterogeneity may improve coalitions' effectiveness. Some papers support the idea that the introduction of heterogeneity yields larger stable coalitions, with or without transfers, while some others find that transfers are necessary to induce larger stable coalitions.

The present paper derives analytical results and proves that introducing heterogeneity in environmental damages does not increase the size of the coalition. On the contrary, if heterogeneity is strong enough, a smaller stable coalition results relative to the homogeneous case. The main difference between our model and those developed by Fuentes-Albero and Rubio (2010) and Pavlova and Zeeuw (2013) is the functional form of the environmental damages, since they use a linear damage function[3]. With a quadratic environmental damage function the analysis becomes more complex but also more interesting, since we can capture the interaction between heterogeneous countries due to the aggregate global emission level. Our results demonstrate that, introducing asymmetry into a stable under symmetry agreement can disturb stability. Moreover, when stable coalitions exist their size is small and, when the asymmetry is strong enough, they can not include both types of countries. Our analysis also confirms that the symmetric approach is a special case of the asymmetric approach. When we simplify the asymmetric analysis, assuming that there exist only one type of countries, the results from our model can be paralleled with those in Rubio and Casino (2001).

The rest of the paper is structured as follows. Section 2 describes the model for the k asymmetric types and solves for the countries' choice of emissions. Section 3 presents the stability conditions. Section 4, studies the two-type case, examines the existence and stability of an IEA when countries are asymmetric in environmental

[3] We employ the damage functional form, $D_i^j(E) = \frac{1}{2}c^j E^2$. To be consistent with the analysis derived in Fuentes-Albero and Rubio (2010) and Pavlova and Zeeuw (2013) our damage function should be simplified to $D_i^j(E) = c^j E$.

damages and presents a counterexample where a stable coalition is not possible. Section 5 concludes the paper.

2 The model

We assume that there are k asymmetric types of countries. Let $K = \{1, ..., k\}$ denote the set of types and the letters $m, j \in K$ denote types. For each type $j \in K$ there exists a set of N^j countries, $N^j = \{1, 2, 3, \ldots, n^j\}$. Let the total set of countries be $N = \bigcup_{j \in K} N^j$ and the total number of countries be $n = \sum_{j \in K} n^j$.

Each country i of type $j \in K$ generates emissions $e_i^j > 0$[4] as a result of its economic activity. It derives benefits, expressed as a function of those emissions $B_i^j(e_i^j)$, which are assumed to be strictly concave, $B_i^j(0) = 0$, $B_i^{j\prime} \geq 0$ and $B_i^{j\prime\prime} < 0$. It also suffers damages from the aggregate emissions of the global pollutant, $D_i^j(E)$, which are assumed to be strictly convex, $D_i^j(0) = 0$, $D_i^{j\prime} \geq 0$ and $D_i^{j\prime\prime} > 0$. In particular and in accordance with the literature, we use the following functional forms,

$$B_i^j(e_i^j) = b^j(a^j e_i^j - \frac{1}{2}(e_i^j)^2) \text{ and } D_i^j(E) = \frac{1}{2}c^j E^2, \tag{1}$$

where a^j, b^j and c^j are type specific, positive parameters, and $E = \sum_{i \in N^j, j \in K} e_i^j$ is the aggregate emission level.

The social welfare of each country i of type j, W_i^j, is defined as the difference between total benefits from its own emissions and environmental damages from aggregate emissions,

$$W_i^j = B_i^j\left(e_i^j\right) - D_i^j(E). \tag{2}$$

Substituting the specific functional forms, country i's of type j social welfare is,

$$W_i^j = b^j\left(a^j e_i^j - \frac{1}{2}\left(e_i^j\right)^2\right) - \frac{1}{2}c^j\left(\sum_{i \in N^j, j \in K} e_i^j\right)^2, \tag{3}$$

where $i \in N^j = \{1, 2, 3, \ldots, n^j\}$ and $j \in K$.

[4]The superscript j denotes the type of the country and the subscript i denotes a particular country belonging to type j.

2.1 Coalition formation

We model the process of the heterogeneous countries' decision as a non-cooperative two stage game and we examine the existence and stability of a self-enforcing coalition aiming at controlling emissions. In the first stage, each country i of type j decides whether or not to join the coalition, while in the second stage, chooses its emission level. We assume that only a single coalition can be formed and we determine the equilibrium number and type of countries participating in the coalition by applying the notions of internal and external stability of a coalition as was originally developed by D'Aspremont et. al (1983) and extended to IEAs by Carraro and Siniscalco (1993) and Barrett (1994). We also assume that when a country contemplates joining or defecting from the coalition, it assumes that no other country will change its decision regarding participation in the coalition. Furthermore, we consider that members of the coalition act cooperatively, maximizing their joint welfare, while non-members act in a non-cooperative way, maximizing their individual welfare, and that in the second stage all countries decide their emission level simultaneously (Cournot approach).

In particular, for each type $j \in K$ a set of countries $S^j \subset N^j$ sign an agreement to reduce the emissions of the global pollutant and $N^j \backslash S^j$ do not. Let $s^j = |S^j|$ for all $j \in K$. Each signatory of type j emits e_s^j, such that $E_{s^j} = s^j e_s^j$, and thus the coalition's total emissions are $E_s = \sum_{j \in K} s^j e_s^j$. Similarly, each non-signatory of type j emits e_{ns}^j, such that $E_{ns^j} = (n^j - s^j) e_{ns}^j$, yielding aggregate emissions of non-signatories, $\sum_{j \in K} (n^j - s^j) e_{ns}^j$. Therefore, the aggregate emission level is, $E = E_s + E_{ns}$, hence, $E = \sum_{j \in K} s^j e_s^j + \sum_{j \in K} (n^j - s^j) e_{ns}^j$, for all $j \in K$.

2.2 Choice of emissions

Signatories maximize the coalition's welfare given by $W_s = \sum_{j \in K} s^j W_s^j$. Therefore, signatories choose e_s^j by solving the following maximization problem,

$$\max_{e_s^j} \sum_{j \in K} s^j \left(B_s^j(e_s^j) - D_s^j(E) \right). \tag{4}$$

Non-signatories maximize their own welfare given by W_{ns}^j by choosing e_{ns}^j. So that,

$$\max_{e_{ns}^j} B_{ns}^j(e_{ns}^j) - D_{ns}^j(E). \tag{5}$$

For each type $j \in K$ let the parameter γ^j be the ratio between environmental damages and benefits due to emissions for all countries N^j. Thus,

$$\gamma^j = \frac{c^j}{b^j}. \tag{6}$$

Let,

$$\Psi = 1 + \sum_{j \in K} \gamma^j (n^j - s^j) + \sum_{j \in K} \gamma^j (s^j)^2 + \sum_{j \in K} \left(\frac{s^j}{b^j} \left(\sum_{m \neq j \in K} c^m s^m \right) \right). \tag{7}$$

The expression Ψ is always positive since $n^j \geq s^j$ and is not type specific. The value of the parameter depends only on the total number of the asymmetric types.

The equilibrium emission level for some signatory country of type $m \in K$ is,

$$e_s^m = a^m - \frac{1}{b^m} \frac{\left(\sum_{j \in K} a^j n^j \right) \left(\sum_{j \in K} c^j s^j \right)}{\Psi}. \tag{8}$$

The aggregate emission level by all signatories is,

$$E_s = \sum_{j \in K} s^j a^j - \frac{\left(\sum_{j \in K} a^j n^j \right)}{\Psi} \sum_{j \in K} \left(\frac{s^j}{b^j} \left(\sum_{j \in K} c^j s^j \right) \right). \tag{9}$$

The equilibrium emission level for some non-signatory country of type $m \in K$ is,

$$e_{ns}^m = a^m - \gamma^m \frac{\left(\sum_{j \in K} a^j n^j \right)}{\Psi}. \tag{10}$$

The aggregate emission level by all non-signatories is,

$$E_{ns} = \sum_{j \in K} (n^j - s^j) a^j - \frac{\left(\sum_{j \in K} a^j n^j \right)}{\Psi} \sum_{j \in K} (n^j - s^j) \gamma^j. \tag{11}$$

The aggregate emission level is, $E = E_s + E_{ns}$, hence,

$$E = \frac{\sum_{j \in K} a^j n^j}{\Psi}. \tag{12}$$

The indirect welfare function for some signatory country of type $m \in K$ is,

$$\mathcal{W}_s^m = \frac{1}{2} b^m \left[(a^m)^2 - \gamma^m \frac{\left(\sum_{j\in K} a^j n^j\right)^2}{\Psi^2} \left(1 + \frac{1}{c^m b^m} \left(\sum_{j\in K} c^j s^j \right)^2 \right) \right]. \qquad (13)$$

The indirect welfare function for some non-signatory country of type $m \in K$ is,

$$\mathcal{W}_{ns}^m = \frac{1}{2} b^m \left[(a^m)^2 - \gamma^m \frac{\left(\sum_{j\in K} a^j n^j\right)^2}{\Psi^2} (1 + \gamma^m) \right]. \qquad (14)$$

3 Stable coalition

To determine the existence and stability of a coalition, we use the notions of the internal and external stability developed by D'Aspremont et. al (1983). The internal stability implies that no coalition member has an incentive to leave the coalition, while the external stability implies that no country outside the coalition has an incentive to join the coalition. In our case, these conditions should be specified for all types of countries, $j \in K$. Let $\mathbf{s}$ be a $k-$dimensional vector that denotes the numbers of signatories of each type, i.e., $\mathbf{s} = (s^1, ..., s^k)$. Similarly, let $\mathbf{s}^{-j}$ be a $k-1$ dimensional vector that denotes the numbers of signatories of all types but j.

Thus, for some country of type $j \in K$, the internal and external stability conditions take the following form respectively:

$$\mathcal{W}_s^j(s^j, \mathbf{s}^{-j}) \geq \mathcal{W}_{ns}^j(s^j - 1, \mathbf{s}^{-j}), \qquad (15)$$

$$\mathcal{W}_s^j(s^j + 1, \mathbf{s}^{-j}) \leq \mathcal{W}_{ns}^j(s^j, \mathbf{s}^{-j}). \qquad (16)$$

In this context, a coalition is characterized stable only if the internal and external conditions are satisfied at the equilibrium $\mathbf{s}$ for all countries of all k types.

Substituting the values of the indirect welfare functions from (13) and (14), the internal and external stability conditions are derived.

The internal stability condition for some country of type $m \in K$ is the following,

$$\frac{1}{2}\gamma^m b^m \left(\sum_{j\in K} a^j n^j\right)^2$$
$$\left[\frac{1+\gamma^m}{\left(\Psi + 2\gamma^m - \frac{1}{b^m}\sum_{j\in K} c^j s^j - c^m \sum_{j\in K}\frac{s^j}{b^j}\right)^2} - \frac{1+\gamma^m\left(\frac{1}{c^m}\sum_{j\in K} c^j s^j\right)^2}{\Psi^2}\right] \geq 0. \tag{17}$$

The external stability condition for some country of type $m \in K$ is the following,

$$\frac{1}{2}\gamma^m b^m \left(\sum_{j\in K} a^j n^j\right)^2 \left[\frac{1+\gamma^m(1+\frac{1}{c^m}\sum_{j\in K} c^j s^j)^2}{\left(\Psi + \frac{1}{b^m}\sum_{j\in K} c^j s^j + c^m \sum_{j\in K}\frac{s^j}{b^j}\right)^2} - \frac{1+\gamma^m}{\Psi^2}\right] \geq 0. \tag{18}$$

4 Two-type case

Considering two types of countries, such that $j \in \{A, B\}$[5], the analysis presented in the general case of k types can be simplified as follows.

For each type $j \in \{A, B\}$ we set the parameters,

$$c = \frac{c^A}{c^B} \text{ and } b = \frac{b^A}{b^B}, \tag{19}$$

where c is the ratio of the slopes of the marginal environmental damages and b is the ratio of the slopes of the marginal benefits, of type A over type B countries.

The expression Ψ takes the form,

$$\Psi = 1 + \gamma^A(n^A - s^A) + \gamma^B(n^B - s^B) + \gamma^A(s^A)^2 + \gamma^B(s^B)^2 + (\frac{c^B}{b^A} + \frac{c^A}{b^B})s^A s^B. \tag{20}$$

Alternative, we can write the expression Ψ as,

$$\Psi = 1+\gamma^A(n^A - s^A)+\gamma^B(n^B - s^B)+\gamma^A(s^A)^2+\gamma^B(s^B)^2+(\gamma^A c^{-1}+\gamma^B c)s^A s^B. \tag{21}$$

[5] For the case with two types, we use the notation $j \in \{A, B\}$ instead of $j \in \{1, 2\}$ for presentation reasons in order to prevent superscript from being interpreted as a power.

Signatories maximize the coalition's welfare given by $W_s = \sum_j s^j W_s^j$, with $j \in \{A, B\}$, that is, $W_s = s^A W_s^A + s^B W_s^B$. Therefore, signatories choose e_s^j by solving the following maximization problem,

$$\max_{e_s^j} \left[s^A \left(B_s^A(e_s^A) - D_s^A(E) \right) + s^B \left(B_s^B(e_s^B) - D_s^B(E) \right) \right], \tag{22}$$

where $E = s^A e_s^A + s^B e_s^B + (n^A - s^A) e_{ns}^A + (n^B - s^B) e_{ns}^B$.

The equilibrium emission levels are,

$$e_s^A = a^A - \frac{\gamma^A (a^A n^A + a^B n^B)(s^A + c^{-1} s^B)}{\Psi}, \tag{23}$$

$$e_s^B = a^B - \frac{\gamma^B (a^A n^A + a^B n^B)(c s^A + s^B)}{\Psi}. \tag{24}$$

The aggregate emission level by all signatories is,

$$E_s = a^A s^A + a^B s^B - \frac{(a^A n^A + a^B n^B)}{\Psi} (s^A + c^{-1} s^B) \left(\gamma^A s^A + c \gamma^B s^B \right). \tag{25}$$

Non-signatories maximize their own welfare given by $W_{ns,}^j$ with $j \in \{A, B\}$, by choosing e_{ns}^j. That is,

$$\max_{e_{ns}^j} \left[B_{ns}^j(e_{ns}^j) - D_{ns}^j(E) \right], \tag{26}$$

where $E = s^A e_s^A + s^B e_s^B + (n^A - s^A) e_{ns}^A + (n^B - s^B) e_{ns}^B$.

The equilibrium emission levels are,

$$e_{ns}^A = a^A - \frac{\gamma^A (a^A n^A + a^B n^B)}{\Psi}, \tag{27}$$

$$e_{ns}^B = a^B - \frac{\gamma^B (a^A n^A + a^B n^B)}{\Psi}. \tag{28}$$

The aggregate emission level by all non-signatories is,

$$E_{ns} = a^A (n^A - s^A) + a^B (n^B - s^B) - \frac{(a^A n^A + a^B n^B)}{\Psi} (\gamma^A (n^A - s^A) + \gamma^B (n^B - s^B)). \tag{29}$$

From (25) and (29), the aggregate emission level is,

$$E = \frac{(a^A n^A + a^B n^B)}{\Psi}. \tag{30}$$

Substituting the equilibrium values of the choice variables from (23), (24), (27) and (28) into equation (3), we derive the indirect welfare function of signatories ($\mathcal{W}_s^A$ and $\mathcal{W}_s^B$) and non-signatories ($\mathcal{W}_{ns}^A$ and $\mathcal{W}_{ns}^B$) for both types of countries.

The indirect welfare functions of signatories are,

$$\mathcal{W}_s^A = \frac{1}{2}b^A\left[(a^A)^2 - \frac{\gamma^A(a^An^A + a^Bn^B)^2(1+\gamma^A(s^A + c^{-1}s^B)^2)}{\Psi^2}\right], \quad (31)$$

$$\mathcal{W}_s^B = \frac{1}{2}b^B\left[(a^B)^2 - \frac{\gamma^B(a^An^A + a^Bn^B)^2(1+\gamma^B(cs^A + s^B)^2)}{\Psi^2}\right]. \quad (32)$$

The indirect welfare functions of non-signatories are,

$$\mathcal{W}_{ns}^A = \frac{1}{2}\left[b^A(a^A)^2 - \frac{\gamma^A(a^An^A + a^Bn^B)^2(1+\gamma^A)}{\Psi^2}\right], \quad (33)$$

$$\mathcal{W}_{ns}^B = \frac{1}{2}b^B\left[(a^B)^2 - \frac{\gamma^B(a^An^A + a^Bn^B)^2(1+\gamma^B)}{\Psi^2}\right]. \quad (34)$$

4.1 The case of homogeneity

Before we proceed we compare our result to the homogeneous case. When countries are identical, it means that there is only one type of countries. Without loss of generality, we assume that $n^A = n^B = \frac{n}{2}$ and $s^A = s^B = \frac{s}{2}$. Moreover, we simplify the parameters as follows[6]: $a^A = a^B = a^I$, $b^A = b^B = b^I$, and $c^A = c^B = c^I$. Therefore, in the symmetric case, $c = b = 1$ since $c = \frac{c^A}{c^B}$ and $b = \frac{b^A}{b^B}$. In addition, we define $\gamma = \frac{c^I}{b^I}$, which indicates the relationship between environmental damages and benefits due to emissions for all countries $i \in N = \{1, 2, 3, \ldots, n\}$. Emissions of signatories are e_s, and of non-signatories e_{ns}. The welfare of signatories and non-signatories are $\mathcal{W}_s$ and $\mathcal{W}_{ns}$, respectively.

The signatories' equilibrium emission level is,

$$e_s = a^I\left(1 - \frac{\gamma sn}{\Psi}\right), \quad (35)$$

where $\Psi = 1 + \gamma(n - s) + \gamma s^2$.

The aggregate emission level by all signatories is,

$$E_s = a^I s\left(1 - \frac{\gamma sn}{\Psi}\right). \quad (36)$$

[6] The superscript I is used to define that countries are identical.

The non-signatories' equilibrium emission level is,

$$e_{ns} = a^I \left(1 - \frac{\gamma n}{\Psi}\right). \tag{37}$$

The aggregate emission level by all non-signatories is,

$$E_{ns} = a^I (n - s) \left(1 - \frac{\gamma n}{\Psi}\right). \tag{38}$$

From (36) and (38), the aggregate emission level is,

$$E = \frac{a^I n}{\Psi}. \tag{39}$$

Substituting the equilibrium values of the choice variables from (35) and (37) into equation (3), assuming that there is only one type of countries, we derive the indirect welfare functions of both signatories ($\mathcal{W}_s$) and non-signatories ($\mathcal{W}_{ns}$).

$$\mathcal{W}_s = \frac{1}{2} \left(a^I\right)^2 b^I \left[1 - \frac{n^2 \gamma (1 + \gamma s^2)}{\Psi^2}\right], \tag{40}$$

$$\mathcal{W}_{ns} = \frac{1}{2} \left(a^I\right)^2 b^I \left[1 - \frac{n^2 \gamma (1 + \gamma)}{\Psi^2}\right]. \tag{41}$$

By collapsing the results of the previous Section to homogeneous countries, we get the same results derived in Rubio and Casino (2001), noting that we use different notations and a slightly different benefit function[7]. Consequently, as expected, the symmetric country is a special case of the asymmetric assumption.

4.2 Existence and stability of a coalition assuming heterogeneity in environmental damages

In order to derive analytical results we restrict the asymmetry between the two types of countries in the environmental damage function. Given that we have to restrict heterogeneity, the choice of keeping heterogeneity of countries' damages seems more appropriate since the strongest part of countries' strategic interactions is captured, in the model, through global pollution. That is, we assume $c^A \neq c^B$ while $a^A = a^B = a^I$ and $b^A = b^B = b^I$[8]. For simplicity we set $n^A = n^B = n$.

[7] Rubio and Casino (2001) assume that the quadratic benefit function for each country takes the form: $Bi(qi) = aqi - \frac{b}{2} qi^2$ with $a > 0$ and $b > 0$.

[8] Following the same notation as in Section 4.1, the superscript I in parameters a and b, i.e. a^I and b^I, is used to define that countries are identical with respect to benefits.

Furthermore, without any loss of generality, we assume that $c > 1$, implying that $c^A > c^B$ and since $b = \frac{b^A}{b^B} = 1$, $\gamma^A > \gamma^B$. Therefore, in this context, type A countries have a steeper marginal environmental damage function compared to type B countries. Thus, type A countries suffer higher marginal environmental damages at any level of global pollution, which implies that they are more sensitive to environmental pollution.

Under these assumptions and using the internal stability condition (17), we derive the internal stability conditions for the two types of countries:

Type A countries,

$$\frac{\gamma^A b^I (a^I n)^2}{2}\left[\frac{1+\gamma^A}{\left(\Psi - 2\gamma^A(s^A - 1) - \Gamma s^B\right)^2} - \frac{1+\gamma^A\left(s^A + c^{-1}s^B\right)^2}{\Psi^2}\right] \geq 0, \tag{42}$$

where $\Gamma = (\gamma^A + \gamma^B)$, $c = \frac{\gamma^A}{\gamma^B}$ (since $b = 1$) and $\Psi = 1 + \gamma^A(n - s^A) + \gamma^B(n - s^B) + \gamma^A(s^A)^2 + \gamma^B(s^B)^2 + \Gamma s^A s^B$.

Type B countries,

$$\frac{\gamma^B b^I (a^I n)^2}{2}\left[\frac{1+\gamma^B}{\left(\Psi - 2\gamma^B(s^B - 1) - \Gamma s^A\right)^2} - \frac{1+\gamma^B\left(cs^A + s^B\right)^2}{\Psi^2}\right] \geq 0. \tag{43}$$

Similarly, using the external stability condition (18), we derive the external stability conditions for the two types of countries:

Type A countries,

$$\frac{\gamma^A b^I (a^I n)^2}{2}\left[\frac{1+\gamma^A\left(1 + s^A + c^{-1}s^B\right)^2}{\left(\Psi + 2\gamma^A s^A + \Gamma s^B\right)^2} - \frac{1+\gamma^A}{\Psi^2}\right] \geq 0. \tag{44}$$

Type B countries,

$$\frac{\gamma^B b^I (a^I n)^2}{2}\left[\frac{1+\gamma^B\left(1 + cs^A + s^B\right)^2}{\left(\Psi + 2\gamma^B s^B + \Gamma s^A\right)^2} - \frac{1+\gamma^B}{\Psi^2}\right] \geq 0. \tag{45}$$

The following result asserts that no stable coalition can contain more than 2 members of the same type.

Lemma 1 *For all $s^j \geq 3$, the internal stability conditions are violated for all $j \in \{A, B\}$.*

Table 1: Stable agreements

		s^B		
		0	1	2
s^A	0		$(0,1)$	$(0,2)$ condition (48)
	1	$(1,0)$	$(1,1)$ condition (46)	$(1,2)$
	2	$(2,0)$ condition (49)	$(2,1)$	$(2,2)$

Therefore, if a stable coalition exists, it can consist of maximum four members since $s^j < 3$, for all $j \in \{A, B\}$. Table 1, presents the cases along with the appropriate conditions under which stable agreements exist.

Lemma 2 *Only the coalitions* $(s^A = 1, s^B = 1)$, $(s^A = 0, s^B = 2)$ *and* $(s^A = 2, s^B = 0)$ *can be stable for* $\gamma^A > \gamma^B$ *and* $n > 3$.

For $n > 3$ and $\gamma^A > \gamma^B$ only the coalitions along the main diagonal of the Table 1 can support stable agreements. For those coalitions, the internal stability conditions are satisfied under some necessary and sufficient conditions. In particular, we have the following three cases.

Case 1:

The coalition $(s^A = 1, s^B = 1)$ is a stable agreement only if,

$$\gamma^A \leq \frac{1}{2\left(n - 2 + \sqrt{4n^2 - 6n + 3}\right)}. \tag{46}$$

In the specific case where $\gamma^A = \gamma^B$ the model represents the symmetric case. Under symmetry, a coalition consisting of two countries is the unique self-enforcing IEA if and only if,

$$\gamma^A = \gamma^B (= \gamma) \leq \frac{1}{n - 4 + 2\sqrt{n^2 - 3n + 3}}. \tag{47}$$

The derived restriction (47) is identical to the one presented in the literature with symmetric countries (De Cara and Rotillon, 2001; Rubio and Casino, 2001) under which an agreement of size two is stable.

Case 2:
The coalition $(s^A = 0, s^B = 2)$ is a stable agreement only if,

$$\gamma^A \leq \frac{2\sqrt{(1+\gamma^B)(1+4\gamma^B)} - \left(1+(3n-2)\gamma^B\right)}{3n},$$
$$\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{48}$$

Case 3:
The coalition $(s^A = 2, s^B = 0)$ is a stable agreement if and only if,

$$\gamma^A \leq \frac{2\left(2+\sqrt{3+n\left(n\left(1-\gamma^B\right)\left(1-4\gamma^B\right)-3\left(1-2\gamma^B\right)\right)}\right) - \left(1+(3n-2)\gamma^B\right)n}{(n-2)(2+3n)},$$
$$\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{49}$$

The following Proposition summarizes the results of the above analysis.

Proposition 3 *Stable coalitions and membership:*
i) The mixed coalition $(s^A = 1, s^B = 1)$ *is stable only under minimal asymmetry, that is, when countries are almost identical* (c^A *is very close to* c^B*).*
ii) When asymmetry increases, the coalition consists only of one type of countries.
iii) When the coalition $(s^A = 0, s^B = 2)$ *is stable, the coalition* $(s^A = 2, s^B = 0)$ *is stable as well.*
iv) When the mixed coalition is stable, the other two coalitions, $(s^A = 0, s^B = 2)$ *and* $(s^A = 2, s^B = 0)$*, are stable as well.*

4.2.1 Aggregate emissions

According to the above analysis, a stable agreement can exist in three possible ways. That is, Case 1: $(s^A = 1, s^B = 1)$, Case 2: $(s^A = 0, s^B = 2)$, and Case 3: $(s^A = 2, s^B = 0)$. We can now compare the aggregate emissions among these three possible cases.

Case 1: $(s^A = 1, s^B = 1)$
The aggregate emission level is,

$$E = \frac{2an}{1+\Gamma(n+1)}. \tag{50}$$

Case 2: $(s^A = 0, s^B = 2)$

The aggregate emission level is,

$$E = \frac{2an}{1 + \Gamma n + 2\gamma^B}. \tag{51}$$

Case 3: $(s^A = 2, s^B = 0)$

The aggregate emission level is,

$$E = \frac{2an}{1 + \Gamma n + 2\gamma^A}. \tag{52}$$

Since $\gamma^A > \gamma^B$, we can easily verify that global emissions are lower in Case 3 and higher in Case 2. Hence, with a high level of asymmetry such that only the coalition $(s^A = 2, s^B = 0)$ satisfies stability, we can achieve the lower level of global emissions.

Lemma 4 *The constraints presented in Section 4.2 guarantee that emissions of both signatories and non-signatories are always positive.*

4.3 Case of instability under heterogeneity

The literature (De Cara and Rotillon, 2001; Finus and Rundshagen, 2001; Rubio and Casino, 2001) has shown that when countries are symmetric, a coalition consisting of two members is the unique self-enforcing agreement. Nonetheless, when we allow countries to be heterogeneous, the analysis shows that asymmetry can have an inverse effect on stability. The result, presented in the Proposition below, indicates that heterogeneity has negative implications on the scope of cooperation relative to the homogeneous case. Specifically, we demonstrate that introducing asymmetry into a stable, under symmetry, coalition could disturb stability.

Proposition 5 *Assuming heterogeneous countries, a stable agreement where* $\sum_j s^{j*} > 1$ *for some* $j \in \{A, B\}$ *may not exist, unlike the case of homogeneous countries.*

Proof. To prove the above proposition, we provide a numerical counterexample where a non-trivial stable coalition does not exist when we relax the homogeneity assumption. We set the following values of the parameters: $a^I = 10$, $b^I = 6$[9] and

[9] Following the same notation as in Section 4.1, a^I and b^I are used to define that countries are identical with respect to benefits.

$n = 5$ implying that $N = 10$, while $c^A = 0.55$ and $c^B = 0.25$. Using these values, we derive, $\gamma^A = \frac{c^A}{b^I} = 0.091\bar{6}$ and $\gamma^B = \frac{c^B}{b^I} = 0.041\bar{6}$.

Consider first the case that all countries are symmetric (they are all of type B). The condition for the coalition ($s = 2$) to be stable is given in (47). For the numerical example, the stability condition requires that $\gamma \leq 0.0433125$, which is satisfied given that $n = 10$, $b^I = 6$ and $c^I = 0.25$. Therefore, in the case of ten type B countries, a coalition of two countries is stable (in accordance to the literature) and the aggregate emission level is given by equation (39), thus $E = \frac{a^I n}{1+\gamma(n-s)+\gamma s^2} = 66.7$.

We now examine stability in the case of two types of countries. Table 2 presents the stability conditions that fail in each of the possible coalitions. As already noted, only the coalitions along the main diagonal of the Table 1 can support stable agreements under the conditions presented in Section 4.2. However, in all three possible coalitions, ($s^A = 1, s^B = 1$), ($s^A = 0, s^B = 2$) and ($s^A = 2, s^B = 0$), the corresponding internal stability conditions are violated. Consequently, stability can be achieved only under the trivial coalition ($s^A = 1, s^B = 0$), indicating that there is no stable agreement where $\sum_j s^{j*} > 1$ for some $j \in \{A, B\}$.

Table 2: No stable agreement

		s^B		
		0	1	2
s^A	0		$-$	$(0,2)$ $W_s^A(0,2) < W_{ns}^A(0,1)$
	1	$(1,0)$	$(1,1)$ $W_s^j(1,1) < W_{ns}^j(0,1)$	$-$
	2	$(2,0)$ $W_s^A(2,0) < W_{ns}^A(1,0)$	$-$	$-$

We first check the stability conditions for the coalition ($s^A = 2, s^B = 0$), i.e. conditions (49). The second condition, given that $n = 5$, yields the following threshold for the parameter γ^B, $\gamma^B < 0.0433125$. This condition is satisfied since for the values of the parameters of the present example, $\gamma^B = 0.041\bar{6}$.

The first of the conditions in (49), given that $n = 5$ and $\gamma^B = 0.041\bar{6}$, requires that $\gamma^A \leq 0.0463334$. This condition is not satisfied, since for the values of the parameters in our example, $\gamma^A = 0.091\bar{6}$. Therefore, the first of the conditions in (49) is violated, implying that the coalition ($s^A = 2, s^B = 0$) is unstable. Note that, both conditions (46) and (48) are more restrictive for the parameter γ^A relative to (49) (Proof, see Appendix). As a consequence, none of the other two coalitions ($s^A = 1, s^B = 1$) and ($s^A = 0, s^B = 2$) can be stable as well. Thus, stability is achieved only under the trivial coalition ($s^A = 1, s^B = 0$) and the aggregate global emission level is $E = \frac{2a^I n}{1+(\gamma^A+\gamma^B)n} = 60$. ■

Therefore, in the case of symmetric type B countries, a stable agreement of size two is possible. On the contrary, if half of the countries are more sensitive to pollution (higher value of c^j) relative to the other halve of type B countries, a stable agreement is not always possible. The latter result holds when asymmetry is very strong, that is, when parameters c^A and c^B differ significantly.

Note that aggregate emissions in the case of ten symmetric type B countries, two of which form a coalition to reduce their emissions, are $E = 66.7$. In the case of five type A and five type B countries, case that does not allow the formation of any stable coalition, aggregate emissions are $E = 60$. Although this is expected since half of the countries (type A countries), being more sensitive to pollution, emit less than the other half (type B countries), it is worth noting that the existence of stable coalitions is not necessary related to lower global emissions.

Figure 2 illustrates the effect of heterogeneity on the stability in the case of the above numerical counterexample. We set $s^B = 0$ and investigate at which s^A the internal stability condition of type A countries is satisfied. In particular, we plot the indirect welfare functions of type A countries against different coalition size s^A when $s^B = 0$. The welfare for the signatories, i.e. $\mathcal{W}_s^A(s^A, s^B)$, is depicted by the solid line and the welfare for the non-signatories, i.e. $\mathcal{W}_{ns}^A(s^A, s^B)$, is depicted by the dotted line. Moreover, the welfare $\mathcal{W}_{ns}^A(s^A - 1, s^B)$ is depicted by the dashed line and represents the welfare for the non-signatories shifted by one (we use that line to represent graphically the internal stability condition).

As indicated in the figure, when $s^B = 0$ the internal stability condition of type A countries, condition (42), is satisfied only at $s^A = 1$. In particular, at this point ($s^A = 1, s^B = 0$) the internal stability condition is satisfied with equality, i.e. $\mathcal{W}_s^A(1, 0) = \mathcal{W}_{ns}^A(0, 0)$. Obviously, at the point ($s^A = 2, s^B = 0$) the condition is violated since $\mathcal{W}_{ns}^A(1, 0) > \mathcal{W}_s^A(2, 0)$. Hence, the only stable coalition is the trivial coalition ($s^A = 1, s^B = 0$) confirming once more that a stable agreement where $\sum_j s^{j*} > 1$ for some $j \in \{A, B\}$ does not exist.

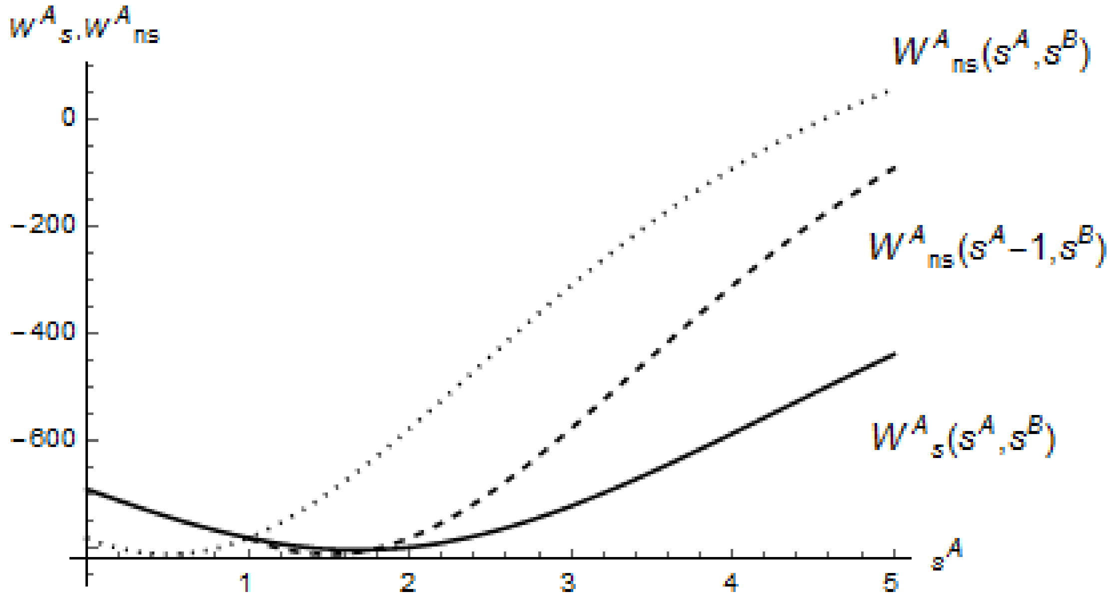

Figure 2

5 Conclusions

The present paper examines the existence and stability of IEAs in a two stage non-cooperative game assuming heterogeneous countries that differ in their sensitivity to the global pollutant. A coalition is considered stable when none of the coalition's members wish to withdraw and no country outside the coalition wishes to join. We use quadratic functions and further assume that in the second stage all countries make their decisions simultaneously.

Our results show that, relaxing the widely used in the literature assumption of symmetric countries, the size of stable coalitions attempting to mitigate environmental problems remains small. The largest possible coalition that can be achieved includes only two countries and the membership of the coalition is mainly driven by the degree of the asymmetry. In particular, the mixed coalition that includes one country of each type, i.e. $(s^A = 1, s^B = 1)$, is possible only when asymmetry is very small. This case is close to the symmetric case, where according to the literature a coalition of two countries is the unique self-enforcing agreement. When heterogeneity is strong enough, a possible coalition consists of two countries again but they belong only to one type, either type A or type B, depending on the level of asymmetry. Under moderate heterogeneity, a coalition can contain either two type B countries, i.e. $(s^A = 0, s^B = 2)$, or two type A countries, i.e. $(s^A = 2, s^B = 0)$. However, when the level of heterogeneity is stronger, a stable coalition can consist only of two type A countries, i.e. $(s^A = 2, s^B = 0)$, and this coalition supports the lower level of global emissions.

An important outcome of the present analysis is that, heterogeneity can have grave implications on the scope of cooperation in comparison with the homogeneous case. We show that, introducing asymmetry into a stable, under symmetry, agreement can disturb stability. We provide a counterexample where a coalition does not exist when countries exhibit a strong level of asymmetry in environmental damages. Consequently, heterogeneity can exacerbate rather than reduce the free-riding incentives.

6 References

Barrett, S. (1994). "Self-enforcing international environmental agreements". *Oxford Economic Papers*, **46**, 878-894.

Barrett, S. (1997). "Heterogenous international environmental agreements". *In: Carraro, C. (Ed.), International Environmental Negotiations*, Edward Elgar, Cheltenham, UK, (Chapter 2).

Biancardi, M. and Villani, G. (2010) "International environmental agreements with asymmetric countries". *Computational Economics*, **36**, 69–92.

Carraro, C. and Siniscalco, D. (1993). "Strategies for the international protection of the environment". *Journal of Public Economics*, **52**, 309-328.

Carraro, C. and Siniscalco, D. (1998). "International environmental agreements - Incentives and political economy". *European Economic Review,* **42**, 561-572.

Chou, P.B. and Sylla, C. (2008). "The formation of an international environmental agreement as a two-stage exclusive cartel formation game with transferable utilities". *International Environmental Agreements*, **8**, 317–341.

D' Aspremont, C., Jacquemin, A., Gabszewicz, J.J., and Weymark, J.A. (1983). "On the stability of collusive price leadership". *Canadian Journal of Economics*, **16**, 17-25.

De Cara, S. and Rotillon, G. (2001). "Multi Greenhouse Gas International Agreements". *Working paper.*

Diamantoudi, E. and Sartzetakis, E. (2006). “Stable International Environmental Agreements: An Analytical Approach”. *Journal of Public Economic Theory, **8***, 247–263.

Finus, M. and Rundshagen, B. (2001). “Endogenous coalition formation global pollution control”. *Working paper*, FEEM, Nota di Lavoro 43.2001.

Fuentes-Alberto, C. and Rubio, S.J. (2010). “Can international environmental cooperation be bought?”. *European Journal of Operational Research*, **202**, 255–264.

Hoel, M. and Schneider, K. (1997). “Incentives to Participate in an International Environmental Agreement”. *Environmental and Resource Economics,* **9**, 153-170.

McGinty, M. (2007). “International environmental agreements among asymmetric nations”. *Oxford Economic Papers*, **59**, 45–62.

Osmani, D. and Tol, R.T.J. (2010). “The case of two self-enforcing international agreements for environmental protection with asymmetric countries”. *Computational Economics*, **36**, 93–119.

Oxfam International (2015).“Extreme Carbon Inequality”. *Series: Oxfam Media Briefings.*

Pavlona, Y. and de Zeeuw, A. (2012). “Asymmetries in International Environmental Agreements”. *Environment and Development Economics*, **18**, 51–68.

Rubio, J. S. and Casino, B. (2001). “International Cooperation in Pollution Control”. *Mimeo.*

Rubio, J. S. and Ulph, A. (2006). “Self-Enforcing International Environmental Agreements Revisited”. *Oxford Economic Papers*, **58**, 233–263.

Sartzetakis, E. and Strantza S. (2013). “International Environmental Agreements: An Emission Choice Model with Abatement Technology”. *Discussion paper*, No. 5/2013.

7 Appendix

In what follows we present the proofs of Lemmas and Propositions in the document.

Proof of Lemma 1. The internal stability condition for type A countries is satisfied if and only if condition (42) is satisfied. Rearranging,

$$\frac{1}{2}\gamma^A b^I (a^I n)^2 \left[\frac{(1+\gamma^A)\Psi^2 - (1+\gamma^A(s^A + c^{-1}s^B)^2)(\Psi - 2\gamma^A(s^A-1) - \Gamma s^B)^2}{\Psi^2(\Psi - 2\gamma^A(s^A-1) - \Gamma s^B)^2}\right] \geq 0. \quad (53)$$

The sign of this condition depends on the sign of the expression in the numerator. Hence,

$$(1+\gamma^A)\Psi^2 - (1+\gamma^A(s^A + c^{-1}s^B)^2)(\Psi - 2\gamma^A(s^A-1) - \Gamma s^B)^2 \geq 0. \quad (54)$$

Recalling that $c = \frac{\gamma^A}{\gamma^B}$, $\Gamma = \gamma^A + \gamma^B$, and $\Psi = 1 + \gamma^A(n - s^A) + \gamma^B(n - s^B) + \gamma^A(s^A)^2 + \gamma^B(s^B)^2 + \Gamma s^A s^B$ and rearranging terms we obtain,

$$(1+\gamma^A)(1 + n\Gamma + \gamma^A s^A(s^A - 1) + \gamma^B s^B(s^B - 1) + \Gamma s^A s^B)^2 - \\ (1 + \gamma^A(s^A + \frac{\gamma^B}{\gamma^A}s^B)^2)(1 + n\Gamma + \gamma^A s^A(s^A - 3) + (\gamma^B s^B - \gamma^A)(s^B - 2) + \Gamma s^A s^B)^2 \geq 0. \quad (55)$$

The term $(1+\gamma^A(s^A+\frac{\gamma^B}{\gamma^A}s^B)^2)$ is greater (or at least equal) to the term $(1+\gamma^A)$ for $s^A \geq 1$. The above expression can be positive only if $s^A < 3$ given s^B. For all $s^A \geq 3$, the second term: $(1 + \gamma^A(s^A + \frac{\gamma^B}{\gamma^A}s^B)^2)(1 + n\Gamma + \gamma^A s^A(s^A - 3) + (\gamma^B s^B - \gamma^A)(s^B - 2) + \Gamma s^A s^B)^2$, is greater than the first term: $(1+\gamma^A)(1+n\Gamma+\gamma^A s^A(s^A - 1) + \gamma^B s^B(s^B - 1) + \Gamma s^A s^B)^2$, and the internal stability condition is violated. The internal stability condition for type B countries is satisfied if and only if condition (43) is satisfied. Rearranging,

$$\frac{1}{2}\gamma^A b^I (an)^2 \left[\frac{(1+\gamma^B)\Psi^2 - (1+\gamma^B(cs^A + s^B)^2)(\Psi - 2\gamma^B(s^B-1) - \Gamma s^A)^2}{\Psi^2(\Psi - 2\gamma^B(s^B-1) - \Gamma s^A)^2}\right] \geq 0. \quad (56)$$

The sign of this condition depends on the sign of the expression in the numerator. Hence,

$$(1+\gamma^B)\Psi^2 - (1+\gamma^B(cs^A + s^B)^2)(\Psi - 2\gamma^B(s^B-1) - \Gamma s^A)^2 \geq 0. \quad (57)$$

Rearranging terms we obtain,

$$(1+\gamma^B)(1 + n\Gamma + \gamma^A s^A(s^A - 1) + \gamma^B s^B(s^B - 1) + \Gamma s^A s^B)^2 - \\ (1 + \gamma^B(\frac{\gamma^A}{\gamma^B}s^A + s^B)^2)(1 + n\Gamma + (\gamma^A s^A - \gamma^B)(s^A - 2) + \gamma^B s^B(s^B - 3) + \Gamma s^A s^B)^2 \geq 0. \quad (58)$$

The term $(1+\gamma^B(\frac{\gamma^A}{\gamma^B}s^A+s^B)^2)$ is greater (or at least equal) to the term $(1+\gamma^B)$ for $s^B \geq 1$. The above expression can be positive only if $s^B < 3$ given s^A. For all $s^B \geq 3$, the second term: $(1+\gamma^B(\frac{\gamma^A}{\gamma^B}s^A+s^B)^2)(1+n\Gamma+(\gamma^A s^A-\gamma^B)(s^A-2)+\gamma^B s^B(s^B-3)+\Gamma s^A s^B)^2$ is greater than the first term: $(1+\gamma^B)(1+n\Gamma+\gamma^A s^A(s^A-1)+\gamma^B s^B(s^B-1)+\Gamma s^A s^B)^2$, and the internal stability condition is violated. Consequently, the two internal stability conditions, i.e. (42) and (43), are satisfied at the equilibrium for all countries of both types $j \in \{A, B\}$ for all $s^j < 3$. ■

Proof of Lemma 2. Table 3, presents more analytically the cases under which stable agreements exist. According to Lemma 1, the two internal stability conditions are satisfied for all $s^j < 3$, for all $j \in \{A, B\}$. Consequently, we have to examine only the cases where $s^j \leq 2$. An agreement is considered stable if the four stability conditions, presented in equations: (42), (43), (44) and (45), are satisfied at the equilibrium. Table 3 includes all the possible coalitions. For each stable coalition, we state the appropriate conditions that ensure stability, while for each non-stable coalition we mention the condition that is violated. For $n > 3$, $\gamma^A > \gamma^B$

Table 3: Possible coalitions

		s^B		
		0	1	2
s^A	0		$(0,1)$ Trivial coalition	$(0,2)$ conditions (68) and (70)
	1	$(1,0)$ Trivial coalition	$(1,1)$ conditions (64) and (66)	$(1,2)$ condition (61)
	2	$(2,0)$ condition (72)	$(2,1)$ condition (62)	$(2,2)$ condition (63)

and $j \in \{A, B\}$ we have the following cases.

Trivial coalition:

Each combination above the main diagonal of Table 3, i.e. $(s^A = 1, s^B = 0)$ and $(s^A = 0, s^B = 1)$, consists a trivial coalition.

The coalitions $(s^A = 1, s^B = 0)$ and $(s^A = 0, s^B = 1)$ are stable if one of the following conditions is satisfied, either,

$$\gamma^B > \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}, \tag{59}$$

or,

$$\gamma^A > \frac{2\left(2+\sqrt{3+n\left(n\left(1-\gamma^B\right)\left(1-4\gamma^B\right)-3\left(1-2\gamma^B\right)\right)}\right)-\left(1+(3n-2)\gamma^B\right)n}{(n-2)(2+3n)},$$

$$\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{60}$$

Violation of internal stability:
The coalitions below the main diagonal of Table 3, i.e. $(s^A = 1, s^B = 2)$, $(s^A = 2, s^B = 1)$ and $(s^A = 2, s^B = 2)$, fail to satisfy the internal stability condition for type B countries. In particular, we have:
At $(s^A = 1, s^B = 2)$ the internal stability condition for type B countries is violated. That is,

$$-\frac{(\gamma^A+3\gamma^B)}{\gamma^B}[(n+1)^2(\gamma^A)^3+(n+1)(3(n+1)\gamma^B+2)(\gamma^A)^2$$
$$+((4+3n)n(\gamma^B)^2+(1+2n)\gamma^B+1)\gamma^A+((n^2-4)(\gamma^B)^2-5\gamma^B-1)\gamma^B]<0. \tag{61}$$

At $(s^A = 2, s^B = 1)$ the internal stability condition for type B countries is violated. That is,

$$-\frac{4(\gamma^A+\gamma^B)}{\gamma^B}[(n+2)^2(\gamma^A)^3+2(n+2)(1+n\gamma^B)(\gamma^A)^2$$
$$+(((n-1)n-3)(\gamma^B)^2+(n-3)\gamma^B+1)\gamma^A-(1+\gamma^B)(1+(n+1)\gamma^B)\gamma^B]<0. \tag{62}$$

At $(s^A = 2, s^B = 2)$ the internal stability condition for type B countries is violated. That is,

$$-\frac{1}{\gamma^B}[4(n+4)(\gamma^A)^4+8(n+4)(2(n+3)\gamma^B+1)(\gamma^A)^3$$
$$+((n+2)(23n+86)(\gamma^B)^2+20(n+3)(\gamma^B)+4))(\gamma^A)^2$$
$$+2((12+7n(n+4))(\gamma^B)^2+(5n-2)(\gamma^B)+2)\gamma^B\gamma^A$$
$$+\left((n-2)(3n+10)(\gamma^B)^2-2(n+14)\gamma^B-5\right)(\gamma^B)^2]<0. \tag{63}$$

Possible stable agreement:
Only the coalitions lying along the main diagonal of Table 3, i.e. $(s^A = 1, s^B = 1)$, $(s^A = 0, s^B = 2)$ and $(s^A = 2, s^B = 0)$, can support stable agreements under some necessary and sufficient conditions.
The coalition $(s^A = 1, s^B = 1)$ is stable under the following conditions:

The internal stability conditions for both types are satisfied. These conditions hold if and only if,

$$\gamma^A \leq \frac{1}{2(n - 2 + \sqrt{4n^2 - 6n + 3})},$$
$$n^2(\gamma^A)^4 + 2n(1 + 2n\gamma^B)(\gamma^A)^3 + ((5n^2 - 2n - 1)(\gamma^B)^2 + (4n - 1)\gamma^B + 1)(\gamma^A)^2 - 2((-n^2 + 2n + 1)\gamma^B + 2)(\gamma^B)^2\gamma^A - (1 + \gamma^B)((2n + 1)\gamma^B + 2)(\gamma^B)^2 \leq 0. \tag{64}$$

The external stability conditions for both types are satisfied. The conditions hold if and only if,

$$(1+\gamma^A(2+\frac{\gamma^B}{\gamma^A})^2)(1+\gamma^A+\gamma^B+n(\gamma^A+\gamma^B))^2-(1+\gamma^A)(1+2(2\gamma^A+\gamma^B)+n(\gamma^A+\gamma^B))^2 \geq 0. \tag{65}$$

Rearranging terms and simplifying,

$$(n^2 - 4)(\gamma^A)^3 + ((4 + 3n)n\gamma^B - 5)(\gamma^A)^2 + ((3(1 + n)^2(\gamma^B)^2 + (2n + 1)\gamma^B) - 1)\gamma^A + ((n + 1)\gamma^B + 1)^2\gamma^B \geq 0. \tag{66}$$

When countries are identical, an agreement consisting of two countries is stable if and only if,

$$\gamma^A = \gamma^B(= \gamma) \leq \frac{1}{n - 4 + 2\sqrt{n^2 - 3n + 3}}. \tag{67}$$

The condition (67) is derived by replacing n with $\frac{n}{2}$ in the first condition in (64), since in the symmetric case $n^A = n^B = \frac{n}{2}$ while in the asymmetric case we assume that $n^A = n^B = n$. The derived restriction (67) is identical to the one presented in the literature with symmetric countries (De Cara and Rotillon, 2001; Rubio and Casino, 2001) under which an agreement of size two is stable.
The coalition $(s^A = 0, s^B = 2)$ is stable under the following conditions:

The internal stability condition for type B countries is satisfied. The condition holds if and only if,

$$\gamma^A \leq \frac{2\sqrt{(1+\gamma^B)(1+4\gamma^B)} - \left(1+(3n-2)\,\gamma^B\right)}{3n},$$
$$\gamma^B < \frac{1}{2(n-2+\sqrt{4n^2-6n+3})}. \tag{68}$$

The external stability conditions for both types are satisfied. The conditions hold if and only if,

$$(1+\gamma^A(1+\frac{2\gamma^B}{\gamma^A})^2)(1+2\gamma^B+n(\gamma^A+\gamma^B))^2-(1+\gamma^A)(1+2(\gamma^A+2\gamma^B)+n(\gamma^A+\gamma^B))^2 \geq 0. \tag{69}$$

Rearranging terms and simplifying,

$$-(n+1)(\gamma^A)^3-((3+n(1-n))\gamma^B+n+2)(\gamma^A)^2+$$
$$(2n(n+2)(\gamma^B)^2+(n-3)\gamma^B-1)\gamma^A+((n+2)\gamma^B+1)^2\gamma^B \geq 0. \tag{70}$$

The external stability conditions (66) and (70) are non-binding if parameter γ^B satisfies the following condition,

$$\gamma^B \geq \frac{2}{5(n-1)+4\sqrt{4n^2-6n+3}}. \tag{71}$$

The coalition ($s^A = 2, s^B = 0$) is stable under the following conditions:

The internal stability condition for type A countries is satisfied. The condition holds if and only if,

$$\gamma^A \leq \frac{2\left(2+\sqrt{3+n\,(n\,(1-\gamma^B)\,(1-4\gamma^B)-3\,(1-2\gamma^B))}\right)-\left(1+(3n-2)\gamma^B\right)n}{(n-2)(2+3n)},$$
$$\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{72}$$

The external stability conditions for both types are satisfied. ■

Proof of Propotision 3. For $\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}$, condition (64), regarding parameter γ^A, is always stricter than condition (68), which is always stricter than condition (72). That is,

$$\frac{2\left(2+\sqrt{3+n\left(n\left(1-\gamma^B\right)\left(1-4\gamma^B\right)-3\left(1-2\gamma^B\right)\right)}\right)-\left(1+(3n-2)\gamma^B\right)n}{(n-2)(2+3n)}>$$
$$\frac{2\sqrt{(1+\gamma^B)(1+4\gamma^B)}-\left(1+(3n-2)\,\gamma^B\right)}{3n}>$$
$$\frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{73}$$

The mixed coalition ($s^A=1, s^B=1$) is stable only if $\gamma^A \leq \frac{1}{2(n-2+\sqrt{4n^2-6n+3})}$. In this case, asymmetry is minimal (γ^A is very close to γ^B implying that c^A is very close to c^B) and countries are almost identical.
When asymmetry increases, meaning that $\gamma^A > \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}$, the mixed coalition is unstable. In this case, a stable coalition consists only of one type of countries and its membership depends on the degree of heterogeneity.
Given that, condition (68), regarding parameter γ^A, is always stricter than condition (72), when the coalition ($s^A=0, s^B=2$) is stable, the coalition ($s^A=2, s^B=0$) is stable as well.
Moreover, given that condition (64), regarding parameter γ^A, is always stricter than condition (68), when the coalition ($s^A=1, s^B=1$) is stable, the coalition ($s^A=0, s^B=2$) is stable as well, and as a consequence the coalition ($s^A=2, s^B=0$) is also stable.
To summarize, when the coalition ($s^A=1, s^B=1$) is stable the other two coalitions, ($s^A=0, s^B=2$) and ($s^A=2, s^B=0$), are stable as well and when the coalition ($s^A=0, s^B=2$) is stable, the coalition ($s^A=2, s^B=0$) is also stable. Therefore, if the coalition ($s^A=2, s^B=0$) fails to satisfy stability requirements, none of the other two coalitions, i.e. ($s^A=0, s^B=2$) and ($s^A=1, s^B=1$), can be stable. ■

Proof of Lemma 4. The emissions of signatories are given by equations (23) and (24). The emissions of non-signatories are given by equations (27) and (28). When countries differ only in environmental damages, emissions are simplified as follows:

$$e_s^A=e_s^B=a^I-\frac{2a^I n(\gamma^A s^A+\gamma^B s^B)}{\Psi}, \tag{74}$$

$$e_{ns}^A=a^I-\frac{2a^I n\gamma^A}{\Psi}, \tag{75}$$

$$e_{ns}^B = a^I - \frac{2a^I n\gamma^B}{\Psi}, \tag{76}$$

where $\Gamma = \gamma^A + \gamma^B$ and $\Psi = 1 + \gamma^A(n - s^A) + \gamma^B(n - s^B) + \gamma^A(s^A)^2 + \gamma^B(s^B)^2 + \Gamma s^A s^B$.

Emissions from both signatories and non-signatories are positive for $n > 3$ and $\gamma^A > \gamma^B$ under the following conditions.

When $(s^A = 0, s^B = 0)$, $(s^A = 1, s^B = 0)$ and $(s^A = 0, s^B = 1)$,

$$\gamma^A \leq \frac{1 + \gamma^B n}{n}. \tag{77}$$

In all other cases,

$$\begin{aligned} \gamma^B &< \frac{1}{(2n - s^A - s^B)(s^A + s^B - 1)}, \\ \gamma^A &\leq \frac{1 + \gamma^B(n - s^B) + \gamma^B s^B(s^B + s^A - 2n)}{(1 + 2n - s^A - s^B)s^A - n}. \end{aligned} \tag{78}$$

For the stable coalitions, i.e. $(s^A = 1, s^B = 1)$, $(s^A = 2, s^B = 0)$ and $(s^A = 0, s^B = 2)$, the condition in (78) regarding parameter γ^B is simplified as follows,

$$\gamma^B < \frac{1}{2(n-1)}. \tag{79}$$

We can verify that for $n > 3$, $2\left(n - 2 + \sqrt{4n^2 - 6n + 3}\right) > 2(n-1)$. Hence, the condition regarding parameter γ^B, i.e. $\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}$, is always stricter than the condition $\gamma^B < \frac{1}{2(n-1)}$. That is,

$$\frac{1}{2\left(n - 2 + \sqrt{4n^2 - 6n + 3}\right)} < \frac{1}{2(n-1)}. \tag{80}$$

Given that $\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}$, the conditions for the parameter γ^A given by (68) and (72) are also stricter than its condition in (78).

Therefore, emissions of both signatories and non-signatories are always positive under any possible stable coalition. ■

NOTE DI LAVORO DELLA FONDAZIONE ENI ENRICO MATTEI
Fondazione Eni Enrico Mattei Working Paper Series

Our Working Papers are available on the Internet at the following addresses:
http://www.feem.it/getpage.aspx?id=73&sez=Publications&padre=20&tab=1

NOTE DI LAVORO PUBLISHED IN 2018

1.2018, CSI Series, Claudio Morana, Giacomo Sbrana, Some Financial Implications of Global Warming: an Empirical Assessment

2.2018, ET Series, Berno Büchel, Stefan Klößner, Martin Lochmüller, Heiko Rauhut, The Strength of Weak Leaders - An Experiment on Social Influence and Social Learning in Teams

3.2018, ET Series, Daniele Valenti, Matteo Manera, Alessandro Sbuelz, Interpreting the Oil Risk Premium: do Oil Price Shocks Matter?

4.2018, CSI Series, Lionel Nesta, Elena Verdolini, Francesco Vona, Threshold Policy Effects and Directed Technical Change in Energy Innovation

5.2018, ET Series, Emerson Melo, A Variational Approach to Network Games

6.2018, ET Series, Daniele Valenti, Modelling the Global Price of Oil: Is there any Role for the Oil Futures-spot Spread?

7.2018, CSI Series, Giovanna d'Adda , Yu Gao , Russell Golman and Massimo Tavoni, It's So Hot in Here : Information Avoidance, Moral Wiggle Room, and High Air Conditioning Usage

8.2018, CSI Series, Cristina Cattaneo, Internal and External Barriers to Energy Efficiency: Made-to-Measure Policy Interventions

9.2018, ET Series, Hipòlit Torró, The Response of European Energy Prices to ECB Monetary Policy

10.2018, ESP Series, Giovanni Occhiali, Giacomo Falchetta, The Changing Role of Natural Gas in Nigeria

11.2018, CSI Series, Nuno Carlos Leitão, Climate Change and Kuznets Curve: Portuguese Experience

12.2018, ET Series, Banchongsan Charoensook, B_i and Branching Strict Nash Networks in Two-way Flow Models: a Generalized Sufficient Condition

13.2018, ET Series, Maryam Ahmadi, Matteo Manera, and Mehdi Sadeghzadeh, Investment-Uncertainty Relationship in the Oil and Gas Industry

14.2018, ET Series, Christian K. Darko, Giovanni Occhiali and Enrico Vanino, The Chinese are Here: Firm Level Analysis of Import Competition and Performance in Sub-Saharan Africa

15.2018, ET Series, Giovanni Gualtieri, Marcella Nicolini, Fabio Sabatini, Luca Zamparelli, Natural Disasters and Demand for Redistribution: Lessons from an Earthquake

16.2018, SAS Series, Linda Arata, Gianni Guastella, Stefano Pareglio, Riccardo Scarpa, Paolo Sckokai, Periurban Agriculture: do the Current EU Agri-environmental Policy Programmes Fit with it?

17.2018, ET Series, HuiHui Liu, ZhongXiang Zhang, ZhanMing Chen, DeSheng Dou, The Impact of China's Electricity Deregulation on Coal and Power Industries: Two-stage Game Modeling Approach

18.2018, ESP Series, ZhongXiang Zhang, Energy Price Reform in China

19.2018, CSI Series, Soheil Shayegh, Valentina Bosetti, Simon Dietz, Johannes Emmerling, Christoph Hambel, Svenn Jensen, Holger Kraft, Massimo Tavoni, Christian Traeger, and Rick Van der Ploeg, Recalculating the Social Cost of Carbon

20.2018, CSI Series, Effrosyni Diamantoudi, Eftichios Sartzetakis, Stefania Strantza, International Environmental Agreements - Stability with Transfers among Countries

21.2018, CSI Series, Branko Bošković, Ujjayant Chakravorty, Martino Pelli, Anna Risch, The Effect of Forest Access on the Market for Fuelwood in India

22.2018, CSI Series, Effrosyni Diamantoudi, Eftichios Sartzetakis, Stefania Strantza, International Environmental Agreements - The Impact of Heterogeneity among Countries on Stability

Fondazione Eni Enrico Mattei
Corso Magenta 63, Milano – Italia

Tel. +39 02.520.36934
Fax. +39.02.520.36946

E-mail: letter@feem.it
www.feem.it

FONDAZIONE ENI
ENRICO MATTEI