International Environmental Agreements - Heterogeneity, Transfers and Issue Linkage

Stefania Strantza

A Thesis
In the Department
of
Economics

Presented in Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy (Economics) at
Concordia University
Montréal, Québec, Canada

May 2018

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Stefania Strantza

Entitled: International Environmental Agreements - Heterogeneity, Transfers and Issue Linkage

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Economics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

Dr. Jochen Jaeger	Chair
Dr. Charles Séguin	External Examiner
Dr. Stylianos Perrakis	External to Program Examiner
Dr. Ming Li	Examiner
Dr. Huan Xie	Examiner
Dr. Effrosyni Diamantoudi	Supervisor

Approved by ______________________
Dr. Damba Lkhagvasuren, Graduate Program Director

June 21, 2018 Dr. André Roy, Dean, Faculty of Arts and Science

ABSTRACT

International Environmental Agreements - Heterogeneity, Transfers and Issue Linkage

Stefania Strantza, Ph.D.
Concordia University, 2018

This thesis is a study of International Environmental Agreements (IEAs). It examines the formation and stability of self-enforcing IEAs by addressing three important aspects: (i) the impact of heterogeneity among countries on stability, (ii) the effect of transfers among heterogeneous countries on stability (iii) the role issue linkage (environmental and trade policies) can play in enhancing international environmental cooperation.

Chapter 1 investigates IEAs among heterogeneous countries in a two-stage non-cooperative game with quadratic benefit and environmental damage functions and simultaneous choice of emissions. The model is solved analytically, and it is shown that stable agreements cannot be larger with asymmetric than symmetric countries. Their size remains small and their membership depends on the degree of heterogeneity. Moreover, results reveal that introducing asymmetry into a stable, under symmetry, agreement can disturb stability. Therefore, the assumption of homogeneity is not the determining factor driving the pessimistic result of very small agreements.

Chapter 2 is an extension of Chapter 1, implementing policies (transfers) that can increase cooperation incentives among heterogeneous countries. It is shown that in the presence of transfers the size of a stable coalition can be enlarged, due to the contribution of those that are more sensitive to environmental pollution. As the degree

of heterogeneity increases, the size of a stable, under transfers, agreement increases as well. However, the analysis demonstrates that reductions in emissions (due to the enlargement of the coalitions) and welfare improvements are rather small, confirming the persistent conclusion in the literature noted as the "paradox of cooperation".

Chapter 3 considers the formation and stability of Global Agreements (GAs). The basic model of the IEAs' literature is extended by letting identical countries choose emission taxes and import tariffs as their policy instruments to manage climate change and control trade. Results illustrate the importance of environmental and trade policies working together to enhance cooperation in effective agreements. Contrary to the IEA model, stable agreements are larger and more efficient in reducing global emissions and improving welfare. Furthermore, the analysis indicates that the size of a stable agreement increases in the number of countries affected by the externalities.

ACKNOWLEDGEMENT

There are many people that have earned my gratitude for their contribution to my time and effort during my Ph.D. studies. First, I would like to express my warmest appreciation to my supervisor Dr. Effrosyni Diamantoudi who made this work possible. Her friendly guidance, motivation, and immense knowledge have been invaluable through all the stages of this thesis.

Moreover, I would like to thank the External Examiner, Dr. Charles Séguin (Department of Economics, Université du Québec à Montréal), the Chair of the Examining Committee, Dr. Jochen Jaeger (Geography, Planning and Environment) and the members of the Examining Committee, Dr. Stylianos Perrakis (Finance), Dr. Ming Li (Economics) and Dr. Huan Xie (Economics). Their insightful comments and suggestions have contributed greatly to the improvement of this thesis.

I am indebted to the Graduate Program Directors, Dr. Szilvia Pápai (former) and Dr. Damba Lkhagvasuren (current), and to the Chairs, Dr. Greg LeBlanc (former) and Dr. Jorgen Hansen (current), of the Department for their great support and guidance during these years. I want to extend my thanks to Dr. Prosper Dovonon (Economics), Dr. Tatyana Koreshkova (Economics) and Dr. Eftichios Sartzetakis (Department of Economics, University of Macedonia) for their help and encouragement. I would like also to acknowledge the former and current administrative staff of the Department (Julia Decker, Lucy Gilson, Lise Gosselin, Bonnie Janicki, Mary-Ann Jirjis and Elise M. Melancon) who were always very kind and supportive.

Last but not least, I want to express my deepest gratitude to my family (Dimitris, Lena, Nikos and Tasos) and friends. This thesis would not have been possible without their warm love, continued patience, and endless support.

Thank you all for encouraging me throughout this experience.

"If you can't explain it simply, you don't understand it well enough."

- Albert Einstein

... to my family, Dimitris, Lena, Nikos

and to you.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

INTRODUCTION

Climate change is arguably the most important and pressing problem humanity currently faces. There is almost unanimous international consensus that "warming of the climate system is unequivocal" and that "human influence on the climate system is clear" (IPCC, 2014). Thus, decisive and speedy policy action to mitigate climate change is required. Although 165 countries have already submitted their pledges to reduce their greenhouse gas emissions, known as the Nationally Determined Contributions (NDCs), following the 21st Conference of the Parties of the United Nations Framework Convention on Climate Change (UNFCCC) in Paris, there are serious doubts as to whether the national pledges will be materialized and, even if they do, whether they will be enough to meet the target of below 2°C increase in average global temperature.

The slow progress of coordinating action at the international level to mitigate climate change is a typical example of the obstacles faced in the provision of public goods (or the mitigation of public bad). Given that costs of reducing Greenhouse Gases (GHGs) are very high while their benefits are spread globally, countries may choose not to implement the necessary policies opting instead to free-ride on other countries' actions. Climate change shares these problems with a number of other global environmental problems, such as ozone depletion, biodiversity and marine pollution. For some of these issues, International Environmental Agreements (IEAs) have been reached, successfully tackling the problem, such as the Montreal Protocol on substances responsible for the depletion of the ozone layer. In some other areas, such as climate change, international negotiations to strengthen actions are still ongoing.

The importance of climate change and the inability of the international community to achieve a global agreement to successfully address the problem, has spurred a substantial literature on IEAs in recent years. A large part of the literature, recognizing the interdependence among countries' choices and the widely spread externalities, which lead to the strategic behavior of countries involved in negotiating IEAs, uses game theory as the tool of analysis. A critical characteristic of IEAs is the lack of a supranational authority that could implement and enforce environmental policies on sovereign states. Like in any other pure public good provision problem, every country has an incentive to free-ride on others' efforts. It does so by avoiding the cost of abating its emissions while at the same time enjoying the benefits of lower aggregate emissions achieved by the countries that remain faithful to the agreement. Since the socially optimal outcome cannot be enforced, IEAs differ from a typical public good and thus, they have to be self-enforcing in the sense that they have to account for the countries' incentives to cheat on or withdraw from the agreement. Consequently, the presence of strong free-riding incentives and the absence of a supranational authority can clearly explain the reasons for which global agreements, for instance Kyoto Protocol, fail to achieve their targets, and so climate change is still speeding.

The main body of the literature studies the formation of IEAs considering that countries signing an agreement (called signatories) form a coalition and maximize the coalition's aggregate welfare, while non-members (called non-signatories) act non-cooperatively maximizing their individual welfare. Within this non-cooperative framework, countries' behavior is modelled as a two-stage game, where in the first stage countries decide whether to join the coalition, while in the second stage they choose their emission level (or abatement level) depending on their membership status. In the second stage, it is assumed that countries either choose emissions (or abatements) simultaneously (Cournot approach) or that the coalition acts as a leader and the non-signatory countries follow (Stackelberg approach). The subgame perfect Nash equilibrium of the resulting two-stage game is usually derived by applying the notions of the internal and external stability conditions developed by D'Aspremont et al. (1983) and extended to IEAs by Carraro and Siniscalco (1993) and Barrett (1994). That is, an IEA is considered to be stable if none of its participating countries has an incentive to withdraw (internal stability) and none of the non-participating countries has an incentive to further join the agreement (external stability), assuming that the remaining players in the game do not revise their membership decision.

Theoretical results are pessimistic. The literature shows that the size of a stable coalition is small, regardless of the total number of countries. The main reason is that strong free-riding incentives discourage countries to take cooperative actions in order to promote environmental protection. 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; 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) 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 therefore per period abatements 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), a stable coalition cannot have more than four members.

The same dismal result is obtained even when the static model is extended to a dynamic framework, which approximates climate change much closer since it introduces stock instead of flow pollutants (Calvo and Rubio, 2013). It is only when the coalition formation is modelled as an infinitely repeated game allowing defectors' punishment that could sustain full cooperation (Barrett, 1999), especially if multiple coalitions are considered (Asheim et al., 2006).

Departing from the assumption of the non-cooperative games, another part of the IEAs' literature applies the core concept of stability to examine coalition formation (Chander and Tulkens, 1995 and 1997). The cooperative approach asserts the formation of the grand coalition and the attainment of efficiency, assuming that when a country deviates it expects that the agreement collapses. The concept of farsighted stability has been used to bridge the gap between these two polar cases. It assumes that when a country defects from an agreement it does not make any assumption regarding the behavior of the coalition's remaining members. Instead, it foresees what their reaction will be, and which equilibrium agreement will result from such a deviation. Diamantoudi and Sartzetakis (2002) formally define the concept of farsighted stability and provide the complete characterization of the farsighted stable set, permitting renegotiation among countries if an IEA collapses. Diamantoudi and Sartzetakis (2015) and (2017) examine respectively the case in which groups of countries may coordinate their actions or act independently, in either joining or withdrawing from an agreement, and in both cases find, using general functional forms, that by not restricting countries to a myopic behavior, increases the set of possible stable coalitions. The above results have been verified in a dynamic setting (de Zeeuw, 2008; Biancardi, 2010) and by using a multi-regional computable general equilibrium model (Lise and Tol, 2004).

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. A number of papers have tried to address the issue by introducing heterogeneity. 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 asymmetries can increase the coalition size. Moreover, Chou and Sylla (2008), Osmani and Tol (2010) and Biancardi and Villani (2010) examine stability considering also two types of countries. In particular, Chou and Sylla (2008) explain theoretically why it is more likely that some developed countries can form a small stable coalition first, and then engage in monetary transfers to form the grand coalition. Osmani and Tol (2010) allow the formation of two separate coalitions and demonstrate that with high environmental damages, forming two coalitions yields higher welfare and better environmental quality relative to a unique coalition. Biancardi and Villani (2010) find that stability depends on the level of the asymmetry and the grand coalition can be obtained only by transfers. Using transfers, Weikard (2009) shows as well that under asymmetry large coalitions may be stable.

In their paper, Fuentes-Albero and Rubio (2010) introduce also two types of countries differing either in abatement costs or environmental damages (which are assumed to be linear on 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 in environmental damages.[1] On the other hand, Pavlova and de 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, and that transfers could improve the gains of cooperation.[2] More recently, Finus and McGinty (2017) employ an abatement choice model introducing any type and degree of asymmetry regarding benefits (which are assumed to be linear in the aggregate contribution) and costs.[3] They show that, under certain conditions, even without transfers the grand

---

[1] Their environmental damage function takes the expression $m_i X$, where $m_i > 0$ is the marginal environmental damage and $X = \sum_{i=1}^{N} x_i$ are the total emissions with x_i capturing the level of emissions generated by country $i \in N$.

[2] In their study, they use an objective function of the form, $\pi_i = \alpha_i(\beta_i q_i - 0.5 q_i^2) - \gamma_i Q$. The first term is the benefit function and the second term is the environmental damage function. Also, q_i denotes emissions by country i, Q is the total emission level and $\alpha_i > 0$, $\beta_i > 0$ and $\gamma_i > 0$.

[3] They apply an abatement model using an objective function of the form, $\pi_i = \alpha_i bQ - \frac{\beta_i c}{2} q_i^2$. The first term is the benefit function from the aggregate abatement level Q and the second term is

coalition can be stable with significant gains from cooperation. The conditions can be relaxed when transfers are introduced.

In the aforementioned papers, transfers are mainly implemented using the optimal transfer scheme. That is, when the coalition payoff equals or exceeds the sum of the outside option payoffs then every coalition member receives at least his free-rider payoff plus a share of the remaining surplus (Eyckmans and Finus, 2004). As the review indicates, results of the theoretical literature are mixed. Some papers support the idea that allowing for heterogeneity yields larger stable coalitions, with and without transfers, while others claim that transfers are necessary to induce cooperation (Petrakis and Xepapadeas, 1996; Botteon and Carraro, 1997 and 2001).

Realizing the importance of heterogeneity, the present thesis modifies the basic (non-cooperative) model of the IEAs' literature with quadratic cost and benefit functions and simultaneous choice of emissions by introducing heterogeneous countries (Chapter 1). The model is solved analytically, and it is shown that allowing for heterogeneity does not yield larger stable agreements. On the contrary, if heterogeneity is strong enough, a smaller stable agreement results relative to the homogeneous case. Chapter 1 complements the existing literature by proving that stable agreements cannot be larger with asymmetric than symmetric countries. Therefore, the assumption of homogeneity is not the determining factor driving the pessimistic result of very small agreements. Heterogeneity can exacerbate rather than reduce the free-riding incentives. To that end, the model is extended by implementing policies

---

the individual country i's abatement cost from its own contribution q_i. Also $\alpha_i > 0$, $\beta_i > 0$, $b > 0$ and $c > 0$.

(transfers) that can increase cooperation incentives among countries (Chapter 2). Results indicate that the presence of transfers can enlarge the size of a stable coalition. However, reductions in emissions (due to the enlargement of the coalitions) and welfare improvements are small. Chapter 2 complements the existing literature by demonstrating that transfers can increase participation in IEAs, but only due to the contribution of those countries that are more sensitive to environmental pollution. As the degree of heterogeneity increases, the size of a stable, under transfers coalition, increases as well. The analysis confirms the "paradox of cooperation" (Barrett, 1994) which states that when cooperation matters most, the associated benefits are minimal, especially when heterogeneity is strong.

Finally, another research line in IEAs' literature, that is related to the present thesis and has to be discussed, is the studies that address the formation of environmental (or climate) coalitions linked to trade policies. During recent years, there has been considerable debate on the extent to which international trade and environmental problems affect each other and whether trade and environmental policies can coherently work together to support effective climate coalitions. Clearly, cooperation in international environmental issues differs from most other international problems (Barrett, 2005). Environment is a global public good, and thus in an IEA members cannot exclude non-members from enjoying the benefits of a better environment. On the other hand, in trade agreements, free trade is not treated as a global public good and non-signatory countries can be excluded from enjoying the benefits of a free trade agreement.

Some of the existing IEAs include provisions that can affect trade. Specifically, trade measures can regulate or restrain the trade in particular materials or products, either between members of the treaty and/or between members and non-members. For example, the Montreal Protocol contains specific trade measures in the form of requirements for a ban on trade between parties and non-parties in products containing or made with ozone-depleting substances (ODS). One of the main goal of those measures was to maximize participation in the Protocol, by excluding non-members from supplies of ODS. According to International Institute for Sustainable Development (IISD), some of the countries that participated in the treaty did so because of the trade provisions. Moreover, trade sanctions have been utilized in the Basel Convention (international transportation of hazardous waste), and the Convention on International Trade in Endangered Species (CITES). With respect to a specific product, the U.S. has applied a penalty tax on foreign automobile manufacturers not meeting the domestic Corporate Average Fuel Economy (CAFE) standards.

The relevant literature of IEA and trade is surprisingly not so extensive. Even though trade measures can be an effective tool in the formation of climate coalitions and should always be taken into consideration when an IEA is designed, only a few papers deal with that issue. Initially, Barrett (1997) analyzes an IEA formation problem in a partial equilibrium model with abatement and illustrates that trade sanctions can help to support cooperation, even full cooperation, among countries. Lessmann, Marschinski, and Edenhofer (2009) apply a dynamic model of climate coalition formation and use trade sanctions as an instrument to promote participation. In their model, coalitions are free to impose tariffs on imports from non-cooperating countries. According to their results, participation in the coalition rises and global welfare also rises along with participation.

More recently, Eichner and Pethig (2013) and (2014) study environmental agreements in a model with international trade. Applying a cap and trade regulation, they find that international trade does not improve the conditions for the formation of effective stable climate coalitions. In particular, agreements are very small and hence ineffective in reducing emissions if coalitions play Nash equilibrium (Eichner and Pethig, 2014), however, agreements may be larger but also ineffective if coalitions behave as Stackelberg leaders (Eichner and Pethig, 2013). Eichner and Pethig (2015), replacing the cap with a tax policy, demonstrate that, when countries employ carbon taxes to fight climate change, the grand coalition is stable (imposing some necessary and sufficient conditions) whether Nash or Stackelberg approaches are assumed.[4] In all the above-mentioned papers (Eichner and Pethig, 2013, 2014 and 2015), the formation of stable self-enforcing IEAs is examined in a free trade world economy. Finally, another study that has explored this idea is the paper by Nordhaus (2015). He applies a numerical general equilibrium model and uses tariff sanctions, that are taken as exogenous, to encourage participation in climate agreements. He finds that trade penalties on non-participants induce a large stable coalition with high levels of abatement.

[4]They use emission taxes as a climate policy instrument while assuming that trade is free among countries.

In the light of the above, it is clear that the interaction between climate coalition formation and trade is very important. Apparently, if non-cooperative countries expect trade costs, they may have incentives to join environmental agreements. The basic assumption underlying the necessity of trade measures is that compliance with an IEA is costly for some countries (i.e. the non-participants) who express strong free-riding incentives. The purpose of trade costs are therefore to prevent those countries (by deteriorating their terms of trade) from enjoying their competitive advantage in trade with other countries controlled by the environmental agreement.

To that end, Chapter 3, extends the basic (non-cooperative) model of the IEAs' literature with symmetric countries, quadratic cost and benefit functions and simultaneous decisions by incorporating trade. To our knowledge, this is the first study in the literature that examines the formation of global agreements where countries choose both emission taxes and import tariffs as their policy instruments to manage climate change and control trade. Results illustrate that the interaction between trade and environment policies is essential to enhance cooperation in effective agreements. Contrary to the IEA model, stable agreements are larger and more efficient in reducing global emissions and improving welfare.

The rest of this thesis proceeds as follows. Chapter 1 examines the formation and stability of self-enforcing IEAs addressing the impact of heterogeneity among countries on stability. Chapter 2 is an extension of Chapter 1 and studies the effect of transfers among heterogeneous countries on stability. Finally, Chapter 3 investigates the role issue linkage (environmental and trade policies) can play in enhancing international environmental cooperation.

CHAPTER 1

INTERNATIONAL ENVIRONMENTAL AGREEMENTS - THE IMPACT OF HETEROGENEITY AMONG COUNTRIES ON STABILITY

1.1 Introduction

The present chapter examines the stability of self-enforcing IEAs among heterogeneous countries in a two-stage, non-cooperative emission game.[1] 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 (Cournot approach). We use quadratic benefit and environmental damage functions and assume k types of countries differing in their sensitivity to the global pollution. The model is solved using backward induction and the subgame perfect Nash equilibrium is derived by applying the notions of the internal and external stability conditions (D'Aspremont et al., 1983).

The main purpose of this study is to examine the formation and stability of environmental agreements relaxing the restrictive and unrealistic assumption of homogeneous countries, an assumption made by many studies in the IEAs' literature.

---

[1] This chapter is a joint work with my supervisor Dr. Effrosyni Diamantoudi (Department of Economics, Concordia University) and Dr. Eftichios Sartzetakis (Department of Economics, University of Macedonia).

It is widely accepted that countries are heterogeneous with respect to costs and benefits due to emissions. In other words, both damages suffered from the global pollutant and benefits derived from emitting the pollutant (related to production and consumption activities) differ significantly among nations.

Our analysis is mostly related to Fuentes-Albero and Rubio (2010) and Pavlova and de Zeeuw (2013) in the sense that we examine self-enforcing IEAs in a model with asymmetric countries. However, our model departs from theirs in that we apply a different functional form regarding environmental damages. In their models, Fuentes-Albero and Rubio (2010) and Pavlova and de Zeeuw (2013) use a damage function that is linear in the aggregate emissions.[2] With a quadratic environmental damage function the analysis becomes more complex but more interesting as well, since we are able to capture the interaction effects between heterogeneous countries due to global pollution (high degree of interdependence among countries).

The main contribution of the present analysis is the proof that, in the absence of policies (e.g. transfers) able to increase cooperation among countries, stable coalitions cannot be larger with asymmetric than symmetric countries. In particular, we provide an analytical solution of the model and show that, in the case of two types, introducing heterogeneity does not enhance the size of a stable agreement compared to the homogeneous case. On the contrary, under heterogeneity, when stable coalitions exist their size remains small and their membership depends on the degree of

[2] 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 de Zeeuw (2013) our damage function should be simplified to $D_i^j(E) = c^j E$.

heterogeneity. That is, the maximum number of countries that cooperate is two and when heterogeneity is strong enough, coalitions cannot include countries belonging to different types. Moreover, results indicate that heterogeneity can reduce the scope of cooperation relative to the homogeneous case. We prove that coalitions that are stable under symmetry they may become unstable when asymmetry is introduced. Therefore, we conclude that the assumption of homogeneity is not the determining factor driving the pessimistic result of very small agreements.

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) who examine the basic model with identical countries, quadratic cost and benefit functions, and simultaneous decisions.[3]

The remaining of the chapter is structured as follows. Section 1.2 describes the model for the k asymmetric types and solves for the countries' choice of emissions. Section 1.3 presents the stability conditions. Section 1.4 studies the two-type case, examines the existence and stability of an IEA when countries are asymmetric in environmental damages and presents a counterexample where a stable coalition is not possible. Section 1.5 concludes.

---

[3] Our model utilizes a slightly different benefit function. Rubio and Casino (2001) assume that the quadratic benefit function for each country takes the form, $B_i(q_i) = aq_i - \frac{b}{2}q_i^2$, where q_i denotes emissions by country i, $a > 0$ and $b > 0$.

1.2 The model

We assume that there are k asymmetric types of countries. Let $K = \{1, 2, 3, ..., k\}$ 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 $n^T = \sum_{j \in K} n^j$.

Each country i of type $j \in K$ generates emissions $e_i^j > 0$ as a result of its economic activity.[4] 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 $B_i^{j\prime\prime} < 0$. Therefore, benefits rise at a decreasing rate. 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 $D_i^{j\prime\prime} > 0$. A convex environmental damage function implies that damages from emissions increase at an increasing rate and so gradually reduce ecosystem services. In other works, more emissions cause more harm on nature.

In particular and in accordance with the literature, we use the following functional forms,

$$B_i^j(e_i^j) = b^j \left(a^j e_i^j - \frac{1}{2}(e_i^j)^2 \right) \text{ and } D_i^j(E) = \frac{1}{2} c^j E^2, \tag{1.1}$$

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 ag-

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

gregate emissions,

$$W_i^j = B_i^j\left(e_i^j\right) - D_i^j(E). \tag{1.2}$$

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

$$W_i^j = b^j\left(a^j e_i^j - \frac{1}{2}\left(e_i^j\right)^2\right) - \frac{1}{2}c^j\left(\sum_{i\in N^j, j\in K} e_i^j\right)^2, \tag{1.3}$$

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

1.2.1 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 (D'Aspremont et al., 1983). 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 $S^j \subset N^j$ signs 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, $E_{ns} = \sum_{j \in K} (n^j - s^j) e_{ns}^j$. 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 $j \in K$.

1.2.2 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,

$$\max_{e_s^j} \sum_{j \in K} \left[s^j \left(B_s^j(e_s^j) - D_s^j(E) \right) \right]. \tag{1.4}$$

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

$$\max_{e_{ns}^j} \left[B_{ns}^j(e_{ns}^j) - D_{ns}^j(E) \right]. \tag{1.5}$$

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,

$$\gamma^j = \frac{c^j}{b^j}. \tag{1.6}$$

Let,

$$\Psi = 1 + \sum_{j \in K} \gamma^j (n^j - s^j) + \sum_{j \in K} \gamma^j (s^j)^2 + \sum_{j \in K} \left(\frac{s^j}{b^j} \left(\sum_{m \neq j \in K} c^m s^m \right) \right). \tag{1.7}$$

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

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

$$e_s^m = a^m - \frac{1}{b^m} \frac{\left(\sum_{j \in K} a^j n^j\right)\left(\sum_{j \in K} c^j s^j\right)}{\Psi}. \tag{1.8}$$

The aggregate emission level by all signatories is,

$$E_s = \sum_{j \in K} s^j a^j - \frac{\left(\sum_{j \in K} a^j n^j\right)}{\Psi} \sum_{j \in K} \left(\frac{s^j}{b^j} \left(\sum_{j \in K} c^j s^j \right) \right). \tag{1.9}$$

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

$$e_{ns}^m = a^m - \gamma^m \frac{\left(\sum_{j \in K} a^j n^j\right)}{\Psi}. \tag{1.10}$$

The aggregate emission level by all non-signatories is,

$$E_{ns} = \sum_{j \in K} (n^j - s^j) a^j - \frac{\left(\sum_{j \in K} a^j n^j\right)}{\Psi} \sum_{j \in K} (n^j - s^j) \gamma^j. \tag{1.11}$$

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

$$E = \frac{\sum_{j \in K} a^j n^j}{\Psi}. \tag{1.12}$$

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

$$\mathcal{W}_s^m = \frac{1}{2} b^m \left[(a^m)^2 - \gamma^m \frac{\left(\sum_{j \in K} a^j n^j\right)^2}{\Psi^2} \left(1 + \frac{1}{c^m b^m} \left(\sum_{j \in K} c^j s^j \right)^2 \right) \right]. \tag{1.13}$$

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

$$\mathcal{W}_{ns}^{m} = \frac{1}{2} b^{m} \left[(a^{m})^{2} - \gamma^{m} \frac{\left(\sum_{j \in K} a^{j} n^{j} \right)^{2}}{\Psi^{2}} (1 + \gamma^{m}) \right]. \tag{1.14}$$

1.3 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, assuming that the remaining countries do not revise their membership decision. 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 forms respectively:

$$\mathcal{W}_{s}^{j}(s^{j}, \mathbf{s}^{-j}) \geq \mathcal{W}_{ns}^{j}(s^{j} - 1, \mathbf{s}^{-j}), \tag{1.15}$$

$$\mathcal{W}_{s}^{j}(s^{j} + 1, \mathbf{s}^{-j}) \leq \mathcal{W}_{ns}^{j}(s^{j}, \mathbf{s}^{-j}). \tag{1.16}$$

In this context, a coalition is characterized stable 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 (1.13) and (1.14), the internal and external stability conditions are derived.

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

$$\frac{1}{2}\gamma^m b^m \left(\sum_{j\in K} a^j n^j\right)^2 \left[\frac{1+\gamma^m}{\left(\Psi + 2\gamma^m - \frac{1}{b^m}\sum_{j\in K} c^j s^j - c^m \sum_{j\in K}\frac{s^j}{b^j}\right)^2} - \frac{1+\gamma^m\left(\frac{1}{c^m}\sum_{j\in K} c^j s^j\right)^2}{\Psi^2}\right] \geq 0. \tag{1.17}$$

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

$$\frac{1}{2}\gamma^m b^m \left(\sum_{j\in K} a^j n^j\right)^2 \left[\frac{1+\gamma^m(1+\frac{1}{c^m}\sum_{j\in K} c^j s^j)^2}{\left(\Psi + \frac{1}{b^m}\sum_{j\in K} c^j s^j + c^m \sum_{j\in K}\frac{s^j}{b^j}\right)^2} - \frac{1+\gamma^m}{\Psi^2}\right] \geq 0. \tag{1.18}$$

1.4 Two-type case

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

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

$$c = \frac{c^A}{c^B} \text{ and } b = \frac{b^A}{b^B}, \tag{1.19}$$

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,

[5]For the two-type case, 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.

$$\Psi = 1 + \gamma^A(n^A - s^A) + \gamma^B(n^B - s^B) + \gamma^A(s^A)^2 + \gamma^B(s^B)^2 + (\frac{c^B}{b^A} + \frac{c^A}{b^B})s^A s^B. \quad (1.20)$$

Alternative, we can write the expression Ψ as,

$$\Psi = 1 + \gamma^A(n^A - s^A) + \gamma^B(n^B - s^B) + \gamma^A(s^A)^2 + \gamma^B(s^B)^2 + (\gamma^A c^{-1} + \gamma^B c)s^A s^B. \quad (1.21)$$

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,

$$\max_{e_s^j} \left[s^A \left(B_s^A(e_s^A) - D_s^A(E) \right) + s^B \left(B_s^B(e_s^B) - D_s^B(E) \right) \right], \quad (1.22)$$

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,

$$e_s^A = a^A - \frac{\gamma^A(a^A n^A + a^B n^B)(s^A + c^{-1}s^B)}{\Psi}, \quad (1.23)$$

$$e_s^B = a^B - \frac{\gamma^B(a^A n^A + a^B n^B)(cs^A + s^B)}{\Psi}. \quad (1.24)$$

The aggregate emission level by all signatories is,

$$E_s = a^A s^A + a^B s^B - \frac{(a^A n^A + a^B n^B)}{\Psi}(s^A + c^{-1}s^B)\left(\gamma^A s^A + c\gamma^B s^B\right). \quad (1.25)$$

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

$$\max_{e_{ns}^j} \left[B_{ns}^j(e_{ns}^j) - D_{ns}^j(E) \right], \tag{1.26}$$

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,

$$e_{ns}^A = a^A - \frac{\gamma^A(a^A n^A + a^B n^B)}{\Psi}, \tag{1.27}$$

$$e_{ns}^B = a^B - \frac{\gamma^B(a^A n^A + a^B n^B)}{\Psi}. \tag{1.28}$$

The aggregate emission level by all non-signatories is,

$$E_{ns} = a^A(n^A - s^A) + a^B(n^B - s^B) - \frac{(a^A n^A + a^B n^B)}{\Psi}(\gamma^A(n^A - s^A) + \gamma^B(n^B - s^B)). \tag{1.29}$$

From (1.25) and (1.29), the aggregate emission level is,

$$E = \frac{(a^A n^A + a^B n^B)}{\Psi}. \tag{1.30}$$

Substituting the equilibrium values of the choice variables from (1.23), (1.24), (1.27) and (1.28) into equation (1.3), we derive the indirect welfare functions 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,

$$\mathcal{W}_s^A = \frac{1}{2} b^A \left[(a^A)^2 - \frac{\gamma^A (a^A n^A + a^B n^B)^2 (1 + \gamma^A (s^A + c^{-1} s^B)^2)}{\Psi^2} \right], \tag{1.31}$$

$$\mathcal{W}_s^B = \frac{1}{2} b^B \left[(a^B)^2 - \frac{\gamma^B (a^A n^A + a^B n^B)^2 (1 + \gamma^B (c s^A + s^B)^2)}{\Psi^2} \right]. \tag{1.32}$$

The indirect welfare functions of non-signatories are,

$$\mathcal{W}_{ns}^A = \frac{1}{2} b^A \left[(a^A)^2 - \frac{\gamma^A (a^A n^A + a^B n^B)^2 (1 + \gamma^A)}{\Psi^2} \right], \tag{1.33}$$

$$\mathcal{W}_{ns}^B = \frac{1}{2} b^B \left[(a^B)^2 - \frac{\gamma^B (a^A n^A + a^B n^B)^2 (1 + \gamma^B)}{\Psi^2} \right]. \tag{1.34}$$

1.4.1 The case of homogeneity

Before we proceed we compare our results 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: $a^A = a^B = a^I$, $b^A = b^B = b^I$, and $c^A = c^B = c^I$.[6] 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 are e_{ns}. The welfare of signatories and non-signatories are $\mathcal{W}_s$ and $\mathcal{W}_{ns}$, respectively.

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

The signatories' equilibrium emission level is,

$$e_s = a^I \left(1 - \frac{\gamma s n}{\Psi}\right), \tag{1.35}$$

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

The aggregate emission level by all signatories is,

$$E_s = a^I s \left(1 - \frac{\gamma s n}{\Psi}\right). \tag{1.36}$$

The non-signatories' equilibrium emission level is,

$$e_{ns} = a^I \left(1 - \frac{\gamma n}{\Psi}\right). \tag{1.37}$$

The aggregate emission level by all non-signatories is,

$$E_{ns} = a^I (n - s) \left(1 - \frac{\gamma n}{\Psi}\right). \tag{1.38}$$

From (1.36) and (1.38), the aggregate emission level is,

$$E = \frac{a^I n}{\Psi}. \tag{1.39}$$

Substituting the equilibrium values of the choice variables from (1.35) and (1.37) into equation (1.3), assuming that there is only one type of countries, we derive the indirect welfare functions of both signatories ($\mathcal{W}_s$) and non-signatories ($\mathcal{W}_{ns}$). Hence,

$$\mathcal{W}_s = \frac{1}{2} \left(a^I\right)^2 b^I \left[1 - \frac{n^2 \gamma (1 + \gamma s^2)}{\Psi^2}\right], \tag{1.40}$$

$$\mathcal{W}_{ns} = \frac{1}{2}\left(a^{I}\right)^{2} b^{I}\left[1 - \frac{n^{2}\gamma(1+\gamma)}{\Psi^{2}}\right]. \tag{1.41}$$

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 approach is a special case of the asymmetric approach.

1.4.2 Existence and stability of a coalition

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 $n^A = n^B = n$. Furthermore, without any loss of generality, we assume that $c > 1$, implying that $c^A > c^B$ and since $b = \frac{b^A}{b^B} = 1$, we have $\gamma^A > \gamma^B$. 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.

[7] Rubio and Casino (2001) assume that the quadratic benefit function for each country takes the form: $B_i(q_i) = aq_i - \frac{b}{2}q_i^2$, where q_i denotes emissions by country i, $a > 0$ and $b > 0$. It is trivial to derive the equivalence between the parameters.

[8] Following the same notation as in Section 1.4.1, the superscript I in parameters a and b, i.e. a^I and b^I, is used to denote that countries are identical with respect to benefits.

Under these assumptions and using the internal stability condition (1.17), we derive the internal stability conditions for the two types of countries:

Type A countries,

$$2\gamma^A b^I (a^I n)^2 \left[\frac{1+\gamma^A}{\left(\Psi - 2\gamma^A(s^A-1) - \Gamma s^B\right)^2} - \frac{1+\gamma^A \left(s^A + c^{-1}s^B\right)^2}{\Psi^2} \right] \geq 0, \qquad (1.42)$$

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,

$$2\gamma^B b^I (a^I n)^2 \left[\frac{1+\gamma^B}{\left(\Psi - 2\gamma^B(s^B-1) - \Gamma s^A\right)^2} - \frac{1+\gamma^B \left(cs^A + s^B\right)^2}{\Psi^2} \right] \geq 0. \qquad (1.43)$$

Similarly, using the external stability condition (1.18), we derive the external stability conditions for the two types of countries:

Type A countries,

$$2\gamma^A b^I (a^I n)^2 \left[\frac{1+\gamma^A \left(1 + s^A + c^{-1}s^B\right)^2}{\left(\Psi + 2\gamma^A s^A + \Gamma s^B\right)^2} - \frac{1+\gamma^A}{\Psi^2} \right] \geq 0. \qquad (1.44)$$

Type B countries,

$$2\gamma^B b^I (a^I n)^2 \left[\frac{1+\gamma^B \left(1 + cs^A + s^B\right)^2}{\left(\Psi + 2\gamma^B s^B + \Gamma s^A\right)^2} - \frac{1+\gamma^B}{\Psi^2} \right] \geq 0. \qquad (1.45)$$

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 $j \in \{A, B\}$.*

Therefore, if a stable coalition exists, it can consist of maximum four members since $s^j < 3$, for all $j \in \{A, B\}$. Table 1.1, presents the cases along with the appropriate conditions under which stable agreements exist.

Table 1.1: Stable agreements

		s^B		
		0	1	2
s^A	0	$(0, 0)$	$(0, 1)$	$(0, 2)$ condition (1.48)
	1	$(1, 0)$	$(1, 1)$ condition (1.46)	$(1, 2)$
	2	$(2, 0)$ condition (1.49)	$(2, 1)$	$(2, 2)$

Lemma 2 *Only the non-trivial coalitions $(s^A = 1, s^B = 1)$, $(s^A = 0, s^B = 2)$ and $(s^A = 2, s^B = 0)$ can be stable.*

Only the non-trivial coalitions along the main diagonal of the Table 1.1 can support stable agreements. For those coalitions, the internal stability conditions are satisfied under some necessary and sufficient conditions. In particular, for $n \geq 3$ and $\gamma^A > \gamma^B$, we have the following three cases.

Case 1:

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

$$\gamma^A \leq \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{1.46}$$

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,

$$\gamma^A = \gamma^B (= \gamma) \leq \frac{1}{n-4+2\sqrt{n^2-3n+3}}. \tag{1.47}$$

The derived restriction (1.47) 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,

$$\begin{aligned} \gamma^A &\leq \frac{2\sqrt{(1+\gamma^B)(1+4\gamma^B)} - \left(1+(3n-2)\gamma^B\right)}{3n}, \\ \gamma^B &< \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \end{aligned} \tag{1.48}$$

Case 3:

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

$$\gamma^A \leq \frac{2\left(2+\sqrt{3+n\left(n\left(1-\gamma^B\right)\left(1-4\gamma^B\right)-3\left(1-2\gamma^B\right)\right)}\right)-\left(1+(3n-2)\gamma^B\right)n}{(n-2)(2+3n)},$$
$$\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{1.49}$$

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.

v) A trivial coalition exists when asymmetry becomes very extreme (c^A and c^B differ significantly).

1.4.2.1 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,

$$E = \frac{2a^I n}{1 + \Gamma(n+1)}. \tag{1.50}$$

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

The aggregate emission level is,

$$E = \frac{2a^I n}{1 + \Gamma n + 2\gamma^B}. \tag{1.51}$$

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

The aggregate emission level is,

$$E = \frac{2a^I n}{1 + \Gamma n + 2\gamma^A}. \tag{1.52}$$

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 1.4.2 guarantee that emissions of both signatories and non-signatories are always positive.*

1.4.3 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 can 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: $a^I = 10$, $b^I = 6$ and $n^A = n^B = 5$ (i.e. $n = 5$), while $c^A = 0.55$ and $c^B = 0.25$.[9] 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 (1.47). 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 (1.39), 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 1.2 presents the stability conditions that fail in each of the possible coalitions. As already noted, only the non-trivial coalitions along the main diagonal of the Table 1.1 can support

[9] Following the same notation as in Section 1.4.1, a^I and b^I are used to denote that countries are identical with respect to benefits.

stable agreements under the conditions presented in Section 1.4.2. 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 1.2: No stable agreement

		s^B		
		0	1	2
s^A	0	$(0,0)$	$-$	$(0,2)$ $W_s^B(0,2) < W_{ns}^B(0,1)$
	1	$(1,0)$	$(1,1)$ $W_s^B(1,1) < W_{ns}^B(1,0)$	$-$
	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. condition (1.49). 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 in the present example $\gamma^B = 0.041\bar{6}$.

The first condition in (1.49), 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 $\gamma^A = 0.091\bar{6}$. Therefore, the first condition in (1.49) is violated, implying that the coalition $(s^A = 2, s^B = 0)$ is unstable. Note that, both conditions (1.46) and (1.48) are more restrictive for the parameter γ^A relative to (1.49) (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 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 1.1 illustrates the effect of heterogeneity on 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).

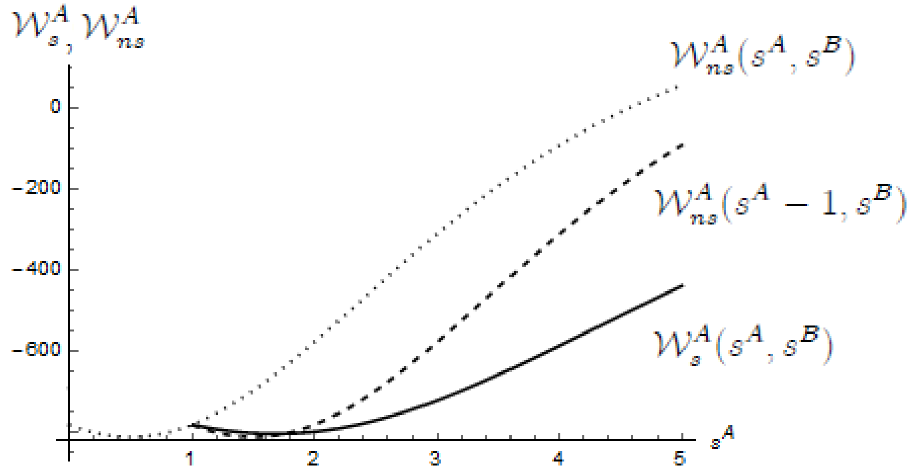


Figure 1.1: Type A countries' welfare functions when $s^B = 0$

As indicated in the figure, when $s^B = 0$ the internal stability condition of type A countries, condition (1.15), 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.

1.5 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 heterogeneity 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.

1.6 Appendix

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

Proof of Lemma 1. The internal stability condition for type A countries is satisfied if and only if condition (1.42) is satisfied. Rearranging,

$$2\gamma^A b^I (a^I n)^2 \left[\frac{(1+\gamma^A)\Psi^2 - (1+\gamma^A(s^A + c^{-1}s^B)^2)(\Psi - 2\gamma^A(s^A-1) - \Gamma s^B)^2}{\Psi^2(\Psi - 2\gamma^A(s^A-1) - \Gamma s^B)^2} \right] \geq 0. \tag{1.53}$$

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

$$(1+\gamma^A)\Psi^2 - (1+\gamma^A(s^A + c^{-1}s^B)^2)(\Psi - 2\gamma^A(s^A-1) - \Gamma s^B)^2 \geq 0. \tag{1.54}$$

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,

$$\begin{gathered}(1+\gamma^A)(1 + n\Gamma + \gamma^A s^A(s^A-1) + \gamma^B s^B(s^B-1) + \Gamma s^A s^B)^2 - \\ (1+\gamma^A(s^A + \frac{\gamma^B}{\gamma^A}s^B)^2)(1 + n\Gamma + \gamma^A s^A(s^A-3) + (\gamma^B s^B - \gamma^A)(s^B-2) + \Gamma s^A s^B)^2 \geq 0.\end{gathered} \tag{1.55}$$

The term $(1+\gamma^A(s^A + \frac{\gamma^B}{\gamma^A}s^B)^2)$ is greater than (or at least equal to) the term $(1+\gamma^A)$ for all $s^A \geq 1$. The above expression can be positive for $s^A < 3$ and $\gamma^A < 2\gamma^B$ 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 (1.43) is satisfied. Rearranging,

$$2\gamma^B b^I(a^I n)^2\left[\frac{(1+\gamma^B)\Psi^2-(1+\gamma^B(cs^A+s^B)^2)(\Psi-2\gamma^B(s^B-1)-\Gamma s^A)^2}{\Psi^2(\Psi-2\gamma^B(s^B-1)-\Gamma s^A)^2}\right]\geq 0. \tag{1.56}$$

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

$$(1+\gamma^B)\Psi^2-(1+\gamma^B(cs^A+s^B)^2)(\Psi-2\gamma^B(s^B-1)-\Gamma s^A)^2\geq 0. \tag{1.57}$$

Rearranging terms we obtain,

$$\begin{aligned}&\left(1+\gamma^B\right)(1+n\Gamma+\gamma^A s^A(s^A-1)+\gamma^B s^B(s^B-1)+\Gamma s^A s^B)^2-\\ &(1+\gamma^B(\frac{\gamma^A}{\gamma^B}s^A+s^B)^2)\left(1+n\Gamma+(\gamma^A s^A-\gamma^B)(s^A-2)+\gamma^B s^B(s^B-3)+\Gamma s^A s^B\right)^2\geq 0.\end{aligned} \tag{1.58}$$

The term $(1+\gamma^B(\frac{\gamma^A}{\gamma^B}s^A+s^B)^2)$ is greater than (or at least equal to) the term $(1+\gamma^B)$ for all $s^B\geq 1$. The above expression can be positive for $s^B<3$ 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 internal stability conditions, i.e. (1.42) and (1.43), 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 1.3 presents analytically the cases under which stable agreements exist. According to Lemma 1, the internal stability conditions are satisfied for all $s^j < 3$, for all $j \in \{A, B\}$. Therefore, we have to examine only the cases where $s^j \leq 2$. An agreement is stable if the stability conditions, presented in equations: (1.42), (1.43), (1.44) and (1.45), are satisfied at the equilibrium. Table 1.3 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 $n \geq 3$ and $\gamma^A > \gamma^B$ we have the following cases.

Table 1.3: Possible coalitions

		s^B		
		0	1	2
s^A	0	$(0,0)$	$(0,1)$ Trivial coalition	$(0,2)$ conditions (1.68) and (1.71)
	1	$(1,0)$ Trivial coalition	$(1,1)$ conditions (1.64) and (1.71)	$(1,2)$ condition (1.61)
	2	$(2,0)$ condition (1.72)	$(2,1)$ condition (1.62)	$(2,2)$ condition (1.63)

Trivial coalition:

Each combination above the main diagonal of Table 1.3, i.e. $(s^A = 1, s^B = 0)$ and $(s^A = 0, s^B = 1)$, consists a trivial coalition. A trivial coalition exists if one of the following conditions is satisfied:

either,

$$\gamma^B > \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}, \tag{1.59}$$

or,

$$\gamma^A > \frac{2\left(2+\sqrt{3+n\left(n\left(1-\gamma^B\right)\left(1-4\gamma^B\right)-3\left(1-2\gamma^B\right)\right)}\right)-\left(1+(3n-2)\gamma^B\right)n}{(n-2)(2+3n)},$$
$$\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{1.60}$$

Violation of internal stability:

The coalitions below the main diagonal of Table 1.3, 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,

$$-\frac{(\gamma^A+3\gamma^B)}{\gamma^B}[(n+1)^2(\gamma^A)^3+(n+1)(3(n+1)\gamma^B+2)(\gamma^A)^2$$
$$+((3n+4)n(\gamma^B)^2+(2n+1)\gamma^B+1)\gamma^A+((n^2-4)(\gamma^B)^2-5\gamma^B-1)\gamma^B]<0. \tag{1.61}$$

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

$$-\frac{4(\gamma^A+\gamma^B)}{\gamma^B}[(n+2)^2(\gamma^A)^3+2(n+2)(1+n\gamma^B)(\gamma^A)^2$$
$$+(((n-1)n-3)(\gamma^B)^2+(n-3)\gamma^B+1)\gamma^A-(1+\gamma^B)(1+(n+1)\gamma^B)\gamma^B]<0. \tag{1.62}$$

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

$$-\frac{1}{\gamma^B}[4(n+4)^2(\gamma^A)^4+8(n+4)(2(n+3)\gamma^B+1)(\gamma^A)^3$$
$$+((n+2)(23n+86)(\gamma^B)^2+20(n+3)(\gamma^B)+4)(\gamma^A)^2$$
$$+2((12+7n(n+4))(\gamma^B)^2+(5n-2)(\gamma^B)+2)\gamma^B\gamma^A$$
$$+\left((n-2)(3n+10)(\gamma^B)^2-2(n+14)\gamma^B-5\right)(\gamma^B)^2]<0. \tag{1.63}$$

Possible stable agreement:

Only the non-trivial coalitions lying along the main diagonal of Table 1.3, 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. The conditions hold if and only if,

$$\gamma^A \leq \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)},$$
$$n^2(\gamma^A)^4+2n(1+2n\gamma^B)(\gamma^A)^3+((5n^2-2n-1)(\gamma^B)^2+(4n-1)\gamma^B+1)(\gamma^A)^2-$$
$$2((-n^2+2n+1)\gamma^B+2)(\gamma^B)^2\gamma^A-(1+\gamma^B)((2n+1)\gamma^B+2)(\gamma^B)^2\leq 0. \tag{1.64}$$

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

$$(1+\gamma^A(2+\frac{\gamma^B}{\gamma^A})^2)(1+\gamma^A+\gamma^B+n(\gamma^A+\gamma^B))^2-(1+\gamma^A)(1+2(2\gamma^A+\gamma^B)+n(\gamma^A+\gamma^B))^2 \geq 0. \tag{1.65}$$

Rearranging terms and simplifying,

$$\begin{aligned}(n^2-4)(\gamma^A)^3+((3n+4)n\gamma^B-5)(\gamma^A)^2+\\ ((3(n+1)^2(\gamma^B)^2+(2n+1)\gamma^B)-1)\gamma^A+((n+1)\gamma^B+1)^2\gamma^B \geq 0.\end{aligned} \tag{1.66}$$

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

$$\gamma^A=\gamma^B(=\gamma) \leq \frac{1}{n-4+2\sqrt{n^2-3n+3}}. \tag{1.67}$$

The condition (1.67) is derived by replacing n with $\frac{n}{2}$ in the first condition in (1.64), 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 (1.67) 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,

$$\gamma^A \leq \frac{2\sqrt{(1+\gamma^B)(1+4\gamma^B)} - \left(1+(3n-2)\,\gamma^B\right)}{3n},$$
$$\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{1.68}$$

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

$$(1+\gamma^A(1+\frac{2\gamma^B}{\gamma^A})^2)(1+2\gamma^B+n(\gamma^A+\gamma^B))^2-(1+\gamma^A)(1+2(\gamma^A+2\gamma^B)+n(\gamma^A+\gamma^B))^2 \geq 0. \tag{1.69}$$

Rearranging terms and simplifying,

$$-(n+1)(\gamma^A)^3 - ((3+n(1-n))\gamma^B + n + 2)(\gamma^A)^2 +$$
$$(2n(n+2)(\gamma^B)^2 + (n-3)\gamma^B - 1)\gamma^A + ((n+2)\gamma^B + 1)^2\gamma^B \geq 0. \tag{1.70}$$

The external stability conditions (1.66) and (1.70) are not binding when parameter γ^B satisfies the following condition,

$$\gamma^B \geq \frac{2}{5(n-1)+4\sqrt{4n^2-6n+3}}. \tag{1.71}$$

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,

$$\gamma^A \leq \frac{2\left(2+\sqrt{3+n\left(n\left(1-\gamma^B\right)\left(1-4\gamma^B\right)-3\left(1-2\gamma^B\right)\right)}\right)-\left(1+(3n-2)\gamma^B\right)n}{(n-2)(2+3n)},$$
$$\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{1.72}$$

The external stability conditions for both types are satisfied. ■

Proof of Proposition 3. For $\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}$, the constraint for the parameter γ^A in (1.64) is always stricter than its constraint in (1.68), which is always stricter than its constraint in (1.72). That is,

$$\frac{2\left(2+\sqrt{3+n\left(n\left(1-\gamma^B\right)\left(1-4\gamma^B\right)-3\left(1-2\gamma^B\right)\right)}\right)-\left(1+(3n-2)\gamma^B\right)n}{(n-2)(2+3n)} >$$
$$\frac{2\sqrt{(1+\gamma^B)(1+4\gamma^B)}-\left(1+(3n-2)\gamma^B\right)}{3n} >$$
$$\frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}. \tag{1.73}$$

The mixed coalition ($s^A = 1, s^B = 1$) is stable only if $\gamma^A \leq \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}$. 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 the constraint for the parameter γ^A in (1.68) is always stricter than its constraint in (1.72), 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 the constraint for the parameter γ^A in (1.64) is always stricter than its constraint in (1.68), 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. Thus, if the coalition ($s^A = 2, s^B = 0$) fails to satisfy stability requirements, none of the other two coalitions, ($s^A = 0, s^B = 2$) and ($s^A = 1, s^B = 1$), can be stable. Therefore, when heterogeneity becomes very extreme (c^A and c^B differ significantly), only a trivial coalition exists. ■

Proof of Lemma 4. The emissions of signatories are given by equations (1.23) and (1.24). The emissions of non-signatories are given by equations (1.27) and (1.28).

When countries differ only in environmental damages, emissions are simplified as follows:

$$e_s^A = e_s^B = a^I - \frac{2a^I n(\gamma^A s^A + \gamma^B s^B)}{\Psi}, \tag{1.74}$$

$$e_{ns}^A = a^I - \frac{2a^I n\gamma^A}{\Psi}, \tag{1.75}$$

$$e_{ns}^B = a^I - \frac{2a^I n\gamma^B}{\Psi}, \tag{1.76}$$

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 of both signatories and non-signatories are positive for $n \geq 3$ and $\gamma^A > \gamma^B$ under the following conditions:

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

$$\gamma^A \leq \frac{1 + \gamma^B n}{n}. \tag{1.77}$$

In all other cases,

$$\gamma^B < \frac{1}{(2n - s^A - s^B)(s^A + s^B - 1)},$$
$$\gamma^A \leq \frac{1 + \gamma^B(n - s^B) + \gamma^B s^B(s^B + s^A - 2n)}{(1 + 2n - s^A - s^B)s^A - n}. \tag{1.78}$$

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 constraint for the parameter γ^B in (1.78) is simplified as follows,

$$\gamma^B < \frac{1}{2\,(n-1)}. \tag{1.79}$$

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

$$\frac{1}{2\left(n - 2 + \sqrt{4n^2 - 6n + 3}\right)} < \frac{1}{2(n-1)}. \tag{1.80}$$

Given that $\gamma^B < \frac{1}{2\left(n-2+\sqrt{4n^2-6n+3}\right)}$, the constraints for the parameter γ^A given by (1.68) and (1.72) are also stricter than its constraint in (1.78).

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

CHAPTER 2

INTERNATIONAL ENVIRONMENTAL AGREEMENTS - STABILITY WITH TRANSFERS AMONG COUNTRIES

2.1 Introduction

The present chapter studies the stability of self-enforcing IEAs among heterogeneous countries allowing for transfers.[1] We employ a two-stage, non-cooperative model, similar to the one presented in Chapter 1, and introduce transfers. The sequence of moves is as follows. In the first stage each country decides whether or not to join the agreement, while in the second stage countries choose their emissions simultaneously. Coalition members agree also, in the first stage, to share the gains from cooperation. We mainly focus on two types of countries differing in their sensitivity to the global pollutant. That is, one type of countries suffers higher environmental damages due to aggregate emissions. In examining the impact of transfers on the coalition size, we apply the notion of Potential Internal Stability (PIS) as defined in Eyckmans and Finus (2004).

It has been shown, in Chapter 1, that introducing heterogeneity does not enhance the size of a stable coalition. Therefore, the main purpose of this study is to examine

[1] This chapter is a joint work with my supervisor Dr. Effrosyni Diamantoudi (Department of Economics, Concordia University) and Dr. Eftichios Sartzetakis (Department of Economics, University of Macedonia).

whether the presence of transfers can successfully increase cooperation incentives among heterogeneous counties and improve the gains of cooperation in terms of reduction in emissions and improvement in welfare levels.

It should be stressed that the main difference between our model and most of the literature is the functional form of the environmental damages. While most papers (Fuentes-Albero and Rubio, 2010; Pavlova and de Zeeuw, 2013) use a linear damage function, we employ a quadratic one.[2] With a quadratic function, we can capture the interaction effects between heterogeneous countries, which seems to have an effect on the results. Thus, in contrast to Pavlova and de Zeeuw (2013), we find that large stable coalitions are possible only with transfers, but when transfers are used we confirm that cooperation requires strong asymmetry. Furthermore, we show that the results obtained by Fuentes-Albero and Rubio (2010) hold also for the case of quadratic environmental damages, but heterogeneity should be stronger to improve cooperation especially when the number of participating countries increases.

In more details, our results indicate that transfers can increase cooperation incentives among heterogeneous countries yielding larger coalition sizes. However, reductions in emissions and thus welfare improvements are small. Furthermore, the inducement of larger coalitions can be achieved only with the help of the countries that suffer the higher damages. That is, stable agreements consist of two, at the maximum, countries of the type with the higher environmental damages and many countries of the type with the lower environmental damages. Strong free-riding in-

[2] Finus and McGinty (2017) employ an abatement choice model and also assume a benefit function that is linear in the aggregate contribution.

centives persist among the type of countries that suffer the higher damages, therefore, only few of them join the coalition. Using transfers, a small number of this type of countries can convince a large number of countries from the other type to join the coalition, but their contribution has small effect on emissions and welfare.

Moreover, the analysis illustrates that as the degree of heterogeneity increases, the size of a stable, under transfers, coalition increases as well. Our findings confirm the persistent conclusion in the literature, first noted in Barrett (1994) and recently noted as the "paradox of cooperation", stating that large stable coalitions do not achieve a lot.

The remaining of this chapter is organized as follows. Section 2.2 describes the model and presents the coalition formation. Section 2.3 solves for the countries' choice of emissions. Section 2.4 analyses the stability of an agreement when countries are heterogeneous in environmental damages and transfers are used to increase cooperation incentives. Section 2.5 presents the aggregate emissions and welfare with and without transfers for stable agreements of different sizes. Section 2.6 concludes.

2.2 The model

We consider two types of countries, $j \in \{A, B\}$. We assume that for each type j there exists a set of N^j countries, $N^j = \{1, 2, 3, \ldots, n^j\}$, each of which generates emissions $e_i^j > 0$ as a result of its economic activity.[3] The set of all countries is defined by N, where $N = N^A \cup N^B$ and the total number of countries is $n^T = n^A + n^B$. Each

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

country i of type j derives benefits from the economic activity, expressed as function of its 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 $B_i^{j\prime\prime} < 0$. Therefore, benefits rise at a decreasing rate. 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 $D_i^{j\prime\prime} > 0$. A convex environmental damage function implies that damages from emissions increase at an increasing rate and so gradually reduce ecosystem services. In other works, more emissions cause more harm on nature.

In particular and in accordance with the literature, we use the following functional forms,

$$B_i^j(e_i^j) = b^j \left(a^j e_i^j - \frac{1}{2}(e_i^j)^2 \right) \text{ and } D_i^j(E) = \frac{1}{2} c^j E^2, \tag{2.1}$$

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

$$E = \sum_{i=1}^{n^A} e_i^A + \sum_{i=1}^{n^B} e_i^B. \tag{2.2}$$

In addition, we incorporate into the model the possibility of welfare transfers T_i^j as well as some form of commitment for those countries that decide to pay the transfers. Transfers T_i^j can be either positive, i.e. $T_i^j > 0$, when a country i of type j receives a payment, or negative, i.e. $T_i^j < 0$, when a country i of type j submits a payment. We make also the standard assumption that transfers balance.

2.2.1 Country's welfare function

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

$$W_i^j = B_i^j \left(e_i^j\right) - D_i^j(E). \tag{2.3}$$

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

$$W_i^j = b^j \left(a^j e_i^j - \frac{1}{2}\left(e_i^j\right)^2\right) - \frac{1}{2}c^j \left(\sum_{i \in N^j, j} e_i^j\right)^2, \tag{2.4}$$

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

2.2.2 Coalition formation

We model the process of the countries' decisions as a non-cooperative two-stage game and 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, emissions are chosen by all countries simultaneously. In addition, in the first stage those countries that decide to join the coalition agree to share the gains from cooperation among its members. Furthermore, we assume that once the agreement is signed, signatories acting as a unique player, maximize the joint welfare, while non-signatories acting in a non-cooperative way, maximize their own welfare. In particular, for each type $j \in \{A, B\}$ a set of countries $S^j \subset N^j$ signs an agreement to reduce the emissions of the global pollutant and the remaining $N^j \backslash S^j$ do not. The game is solved by backward induction.

Once emissions have been chosen and welfare levels have been realized, transfers are implemented.

Following D'Aspremont et al. (1983), we define a stable coalition as a coalition which is both internally and externally stable. Stable agreements are those from which no signatory country has incentives to leave (internal stability) and no country outside the agreement has incentives to join (external stability), assuming that the rest of the countries do not change their membership decision. Thus, the stability conditions, for type A and B countries respectively, take the following forms:

internal stability conditions,

$$\begin{aligned} \mathcal{W}_s^A(s^A, s^B) &\geq \mathcal{W}_{ns}^A(s^A - 1, s^B) \text{ and} \\ \mathcal{W}_s^B(s^A, s^B) &\geq \mathcal{W}_{ns}^B(s^A, s^B - 1), \end{aligned} \tag{2.5}$$

external stability conditions,

$$\begin{aligned} \mathcal{W}_s^A(s^A + 1, s^B) &\leq \mathcal{W}_{ns}^A(s^A, s^B) \text{ and} \\ \mathcal{W}_s^B(s^A, s^B + 1) &\leq \mathcal{W}_{ns}^B(s^A, s^B), \end{aligned} \tag{2.6}$$

where $s^j = |S^j|$ denotes the number of type $j \in \{A, B\}$ countries that sign the agreement, $\mathcal{W}_s^j$ is the welfare of a signatory country and $\mathcal{W}_{ns}^j$ is the welfare of a non-signatory country.

To explore the scope of cooperation when countries use transfers, we apply the Potential Internal Stability (PIS) condition as defined in Eyckmans and Finus (2004). This condition implies that the aggregate net benefits of the coalition must exceed the aggregate of the outside net-benefit options of all coalition members. Hence, countries can redistribute payoffs within the coalition such that the coalition is internally stable. The following condition should be satisfied to ensure that a coalition is potentially internally stable,

$$\sum_{j\in\{A,B\}} s^j \mathcal{W}_s^j(s^j, s^{-j}) \geq \sum_{j\in\{A,B\}} s^j \mathcal{W}_{ns}^j(s^j - 1, s^{-j}). \tag{2.7}$$

That is, the aggregate welfare of all coalition members should be at least larger than the aggregate welfare they receive deciding to free-ride. In other words, the above condition states that the sum of the internal stability conditions should be non-negative.

It follows that the sum of the internal stability conditions in the case of transfers is the sum of the internal stability conditions for the case without transfers, since transfers add up to zero. Recall that, we make the standard assumption that transfers balance, i.e. $\sum_{i\in S^j, j} T_i^j = 0$, where $j \in \{A, B\}$. That is, $\sum_{i=1}^{s^A} T_i^A + \sum_{i=1}^{s^B} T_i^B = s^A T_s^A + s^B T_s^B = 0$. This leads to the following internal stability condition,

$$\begin{aligned} PIS(s^A, s^B) \;\; = \;\; & s^A[\mathcal{W}_s^A(s^A, s^B) - \mathcal{W}_{ns}^A(s^A - 1, s^B)] + \\ & s^B[\mathcal{W}_s^B(s^A, s^B) - \mathcal{W}_{ns}^B(s^A, s^B - 1)] \geq 0. \end{aligned} \tag{2.8}$$

The option of transfers may allow coalition members to allocate their net benefits in such a way that a larger number of countries will have no incentives to leave the coalition. Thus, there could be a self-financed transfer T_i^j from the i cooperating countries of type j to the other non-cooperating countries that can successfully enlarge the original coalition. The potential internal stability is a sufficient condition for internal stability in the presence of transfers, provided that transfers are optimally designed. According to Eyckmans and Finus (2004), under an optimal transfer scheme every coalition member receives at least its free-rider payoff and there may be an extra share of the surplus $PIS(s^A, s^B)$.

When transfers are used to increase cooperation, the stability conditions, for each type of country, are modified as follows:

internal stability conditions,

$$\begin{aligned} \mathcal{W}_s^A(s^A, s^B) + T_s^A(s^A, s^B) &\geq \mathcal{W}_{ns}^A(s^A - 1, s^B) \text{ and} \\ \mathcal{W}_s^B(s^A, s^B) + T_s^B(s^A, s^B) &\geq \mathcal{W}_{ns}^B(s^A, s^B - 1), \end{aligned} \tag{2.9}$$

external stability conditions,

$$\begin{aligned} \mathcal{W}_s^A(s^A + 1, s^B) + T_s^A(s^A + 1, s^B) &\leq \mathcal{W}_{ns}^A(s^A, s^B) \text{ and} \\ \mathcal{W}_s^B(s^A, s^B + 1) + T_s^B(s^A, s^B + 1) &\leq \mathcal{W}_{ns}^B(s^A, s^B). \end{aligned} \tag{2.10}$$

In other words, internal stability holds when the welfare of a signatory country net of the transfer, which could be positive or negative, is larger than its welfare under the free-riding option. On the other hand, external stability holds when a non-signatory country's welfare exceeds the welfare it earns when it is part of the agreement, taking into account the transfer payment.

2.3 Choice of emissions

We solve the game using backward induction. Thus, once emissions have been chosen and welfare levels have been realized, transfers are implemented to examine their effect on the game. Each signatory of type j emits e_s^j, such that $E_{s^j} = s^j e_s^j$, where $s^j = |S^j|$, and thus the coalition's total emissions are $E_s = E_{s^A} + E_{s^B}$. 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 $E_{ns} = E_{ns^A} + E_{ns^B}$. Therefore, global emissions are given by,

$$E = E_s + E_{ns} = s^A e_s^A + s^B e_s^B + (n^A - s^A)e_{ns}^A + (n^B - s^B)e_{ns}^B. \tag{2.11}$$

Before we proceed to the solutions regarding countries' emissions and welfare levels, we define the following parameters in order to simplify the presentation. Namely, parameter γ^j indicates the relationship between environmental damages and benefits due to emissions for all countries i in type $j \in \{A, B\}$. Thus,

$$\gamma^j = \frac{c^j}{b^j}. \tag{2.12}$$

Moreover, parameters c and b are defined as follows,

$$c = \frac{c^A}{c^B} \text{ and } b = \frac{b^A}{b^B}, \tag{2.13}$$

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.

Finally, we define the expression Ψ,

$$\Psi = 1 + \gamma^A(n^A - s^A) + \gamma^B(n^B - s^B) + \gamma^A(s^A)^2 + \gamma^B(s^B)^2 + (\frac{c^B}{b^A} + \frac{c^A}{b^B})s^A s^B, \quad (2.14)$$

which can also be written as,

$$\Psi = 1+\gamma^A(n^A-s^A)+\gamma^B(n^B-s^B)+\gamma^A(s^A)^2+\gamma^B(s^B)^2+(\gamma^A c^{-1}+\gamma^B c)s^A s^B. \quad (2.15)$$

Note that Ψ is always positive since $s^A \leq n^A$, $s^B \leq n^B$ and $\gamma^j > 0$.

The payoff function for each country i of type j, is given by equation (2.4). Each country receives benefits from its economic activity while it suffers damages from global emissions. Signatories maximize the coalition's welfare given by $W_s = \sum_j s^j W_s^j$, where $j \in \{A, B\}$, 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,

$$\max_{e_s^j} \left[s^A \left(B_s^A(e_s^A) - D_s^A(E)\right) + s^B \left(B_s^B(e_s^B) - D_s^B(E)\right)\right], \quad (2.16)$$

where aggregate emissions E are given by equation (2.11).

The first order conditions of the signatories' maximization problem (2.16) yield the equilibrium emissions,

$$e_s^A = a^A - \frac{\gamma^A(a^A n^A + a^B n^B)(s^A + c^{-1}s^B)}{\Psi}, \quad (2.17)$$

$$e_s^B = a^B - \frac{\gamma^B(a^A n^A + a^B n^B)(cs^A + s^B)}{\Psi}. \quad (2.18)$$

Non-signatories choose their emissions e^j_{ns}, by maximizing their own welfare given by W^j_{ns}, where $j \in \{A, B\}$. Hence, they solve the following maximization problem,

$$\max_{e^j_{ns}} \left[B^j_{ns}(e^j_{ns}) - D^j_{ns}(E) \right], \tag{2.19}$$

where aggregate emissions E are given by equation (2.11).

The first order conditions of the non-signatories' maximization problem (2.19) yield the equilibrium emissions,

$$e^A_{ns} = a^A - \frac{\gamma^A(a^A n^A + a^B n^B)}{\Psi}, \tag{2.20}$$

$$e^B_{ns} = a^B - \frac{\gamma^B(a^A n^A + a^B n^B)}{\Psi}. \tag{2.21}$$

Substituting the equilibrium values of the choice variables from (2.17), (2.18), (2.20) and (2.21) into equation (2.11), we derive the aggregate emissions,

$$E = \frac{(a^A n^A + a^B n^B)}{\Psi}. \tag{2.22}$$

We continue by substituting the equilibrium values of the choice variables from (2.17), (2.18), (2.20) and (2.21) into equation (2.4), to derive the indirect welfare functions of signatories ($\mathcal{W}^A_s$ and $\mathcal{W}^B_s$) and non-signatories ($\mathcal{W}^A_{ns}$ and $\mathcal{W}^B_{ns}$) for both types of countries.

The indirect welfare functions of signatories are,

$$\mathcal{W}_s^A = \frac{1}{2}b^A\left[(a^A)^2 - \frac{\gamma^A(a^An^A + a^Bn^B)^2(1+\gamma^A(s^A + c^{-1}s^B)^2)}{\Psi^2}\right], \tag{2.23}$$

$$\mathcal{W}_s^B = \frac{1}{2}b^B\left[(a^B)^2 - \frac{\gamma^B(a^An^A + a^Bn^B)^2(1+\gamma^B(cs^A + s^B)^2)}{\Psi^2}\right]. \tag{2.24}$$

The indirect welfare functions of non-signatories are,

$$\mathcal{W}_{ns}^A = \frac{1}{2}b^A\left[(a^A)^2 - \frac{\gamma^A(a^An^A + a^Bn^B)^2(1+\gamma^A)}{\Psi^2}\right], \tag{2.25}$$

$$\mathcal{W}_{ns}^B = \frac{1}{2}b^B\left[(a^B)^2 - \frac{\gamma^B(a^An^A + a^Bn^B)^2(1+\gamma^B)}{\Psi^2}\right]. \tag{2.26}$$

2.4 Stable coalitions with transfers

Without permitting transfers, it has been shown in Chapter 1 that under heterogeneity the size of stable coalitions remains small and in some cases smaller than in the case of homogeneity. That is, heterogeneity can exacerbate rather than reduce free-riding incentives. In this section, we examine whether transfers can be used to increase participation in an IEA.

We focus on internal stability and recall that potential internal stability is given in condition (2.8). Substituting the values of the indirect welfare functions from (2.23), (2.24), (2.25) and (2.26), into condition (2.8) yields,

$$PIS(s^A, s^B) = \frac{\left(a^A n^A + a^B n^B\right)^2}{2} \left\{ \begin{array}{l} s^A \gamma^A b^A \left[\frac{1+\gamma^A}{\left(\Psi - 2\gamma^A (s^A - 1) - \gamma^A (b + c^{-1}) s^B\right)^2} - \frac{1+\gamma^A \left(s^A + c^{-1} s^B\right)^2}{\Psi^2} \right] \\ + s^B \gamma^B b^B \left[\frac{1+\gamma^B}{\left(\Psi - 2\gamma^B (s^B - 1) - \gamma^B (b^{-1} + c) s^A\right)^2} - \frac{1+\gamma^B \left(c s^A + s^B\right)^2}{\Psi^2} \right] \end{array} \right\} \geq 0, \tag{2.27}$$

where, as previously defined, $b = \frac{b^A}{b^B}$, $c = \frac{c^A}{c^B}$, $\gamma^j = \frac{c^j}{b^j}$ with $j \in \{A, B\}$, and $\Psi = 1 + \gamma^A(n^A - s^A) + \gamma^B(n^B - s^B) + \gamma^A(s^A)^2 + \gamma^B(s^B)^2 + (\gamma^A c^{-1} + \gamma^B c)s^A s^B$.

2.4.1 Heterogeneity in environmental damages

In order to derive analytical results in Chapter 1, we restrict heterogeneity among countries. In order to compare results, we make the same assumption, that is, countries are assumed to be heterogeneous in the environmental damages, while they have the same benefit 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 consider $c^A \neq c^B$ while $a^A = a^B = a^I$ and $b^A = b^B = b^I$.[4] Furthermore, without any loss of generality, we assume that $c > 1$, implying that $c^A > c^B$, and since $b = \frac{b^A}{b^B} = 1$, we have $\gamma^A > \gamma^B$. Therefore, in this context, type A countries have a steeper marginal environmental damage function compared to type B countries. That is, type A countries suffer higher marginal environmental damages at any level of global pollution, which implies that they are more sensitive

[4]The superscript I in parameters a and b, i.e. a^I and b^I, is used to denote that countries are identical with respect to benefits.

to environmental pollution. For simplicity and without any loss of generality we set $n^A = n^B = n$.

Under these assumptions, the PIS condition can be written as follows,

$$PIS(s^A, s^B) = 2b^I(a^I n)^2 \left\{ \begin{array}{l} s^A\gamma^A \left[\frac{1+\gamma^A}{(\Psi - 2\gamma^A(s^A-1) - \Gamma s^B)^2} - \frac{1+\gamma^A\left(s^A + c^{-1}s^B\right)^2}{\Psi^2} \right] \\ +s^B\gamma^B \left[\frac{1+\gamma^B}{(\Psi - 2\gamma^B(s^B-1) - \Gamma s^A)^2} - \frac{1+\gamma^B\left(cs^A + s^B\right)^2}{\Psi^2} \right] \end{array} \right\} \geq 0. \tag{2.28}$$

Moreover, based on the above assumptions regarding the parameters, the expression Ψ can be written as $\Psi = 1+\gamma^A(n-s^A)+\gamma^B(n-s^B)+\gamma^A(s^A)^2+\gamma^B(s^B)^2+\Gamma s^A s^B$, where $\Gamma = \gamma^A + \gamma^B$. Note that $c = \frac{\gamma^A}{\gamma^B}$ since $b = 1$.

Given the higher sensitivity of type A countries ($c > 1$), they benefit from cooperation that yields lower levels of global pollution and are willing to provide side payments to less sensitive type B countries, in order to support a large stable coalition. We are interested in finding the number of type A countries, i.e. s^A, that sign the agreement and agree to share the gains from cooperation as well as the maximum number of type B countries, i.e. s^B, that are lured into signing the agreement by transfers at a level at least equal to their free-riding gains.

Assuming the same type of heterogeneity but without transfers, in the study presented in Chapter 1, we prove analytically that the largest possible stable coalition that can be achieved includes only two countries and the membership of the coalition is mainly driven by the degree of heterogeneity in environmental damages. In particular, there are three possible cases: a mixed coalition that includes one country of each type, i.e. ($s^A = 1$, $s^B = 1$), when heterogeneity is small, a coalition with two type B countries, i.e. $(s^A = 0,\ s^B = 2)$, when heterogeneity is moderate, and a coalition with two type A countries, i.e. $(s^A = 2,\ s^B = 0)$, when heterogeneity is strong. Moreover, when heterogeneity exceeds a certain level, a stable non-trivial coalition does not exist.

Unfortunately, it is not possible to derive analytical results when transfers are introduced. Thus, we resort to simulations. The following Remark summarizes the results obtained using simulations.

Remark 6 *Allowing for transfers among heterogeneous countries increases cooperation. However, the increase in the coalition size does not come from countries belonging to the most suffering type (type A). Only type B countries join the coalition because of the transfers they receive.*

It is evident that the introduction of transfers cannot not induce more type A countries to cooperate, since these countries have to provide the necessary transfers to type B countries. Thus, we can have a coalition with either $s^A = 1$ or $s^A = 2$ type A signatories. For any $s^A \geq 3$ the internal stability condition for type A countries (first condition in (2.9)) is not satisfied. This result was expected since the need to provide transfer payments exacerbates the existing free-riding incentives. The number of type B countries that are willing to join the coalition, under the condition that they receive transfers, varies depending on the degree of heterogeneity.

Our simulations demonstrate that as the degree of heterogeneity increases, the size of the coalition under transfers increases as well. Without loss of generality we set $s^B = n$ assuming that all type B countries participate in the agreement (because of the transfers they receive) and we try to find the degree of heterogeneity required to support coalitions of different sizes, i.e. either ($s^A = 1, s^B = n$) or ($s^A = 2, s^B = n$), for any number of countries $n \geq 3$.

Setting $s^B = n$ and rearranging terms, the PIS condition (2.28) can be written as follows,

$$PIS(s^A, n) = 2b^I(a^I n)^2 \left\{ \begin{array}{c} s^A\gamma^A \left[\frac{1+\gamma^A}{(\Psi - 2\gamma^A(s^A-1) - \Gamma n)^2} - \frac{1+\gamma^A\left(s^A + c^{-1}n\right)^2}{\Psi^2} \right] \\ +n\gamma^B \left[\frac{1+\gamma^B}{(\Psi - 2\gamma^B(n-1) - \Gamma s^A)^2} - \frac{1+\gamma^B\left(cs^A + n\right)^2}{\Psi^2} \right] \end{array} \right\} \geq 0. \tag{2.29}$$

Given the assumption, $s^B = n$, the value of Ψ can be written as $\Psi = 1 + \gamma^A(n - s^A) + \gamma^A(s^A)^2 + \gamma^B n^2 + \Gamma s^A n$.

Condition (2.29) is satisfied, only if $s^A \in \{1, 2\}$ while $s^B = n$ for any value of $n \geq 3$. Moreover, in all cases presented in our numerical analysis, if condition (2.29) holds, meaning that an enlarged coalition is internally stable, the external stability condition for type A countries (first condition in (2.10)) holds as well.

In the following table, Table 2.1, we summarize all possible stable coalitions that can be achieved for $n \in \{3, 4, ..., 10\}$ and $s^A \in \{1, 2\}$ presenting the threshold values of parameters γ^A and γ^B that support each one of them.[5] Note that, as indicated from the expression in (2.29), the relationship between γ^A and γ^B is not linear. When, for instance, parameter γ^B takes its maximum value, condition (2.29) specifies the maximum value parameter γ^A can take so that the corresponding coalition is

[5] Values are rounded off so that they do not exceed their corresponding thresholds.

stable. According to the analysis, a larger coalition requires stricter constraints for the parameters of the model, i.e. γ^A and γ^B.

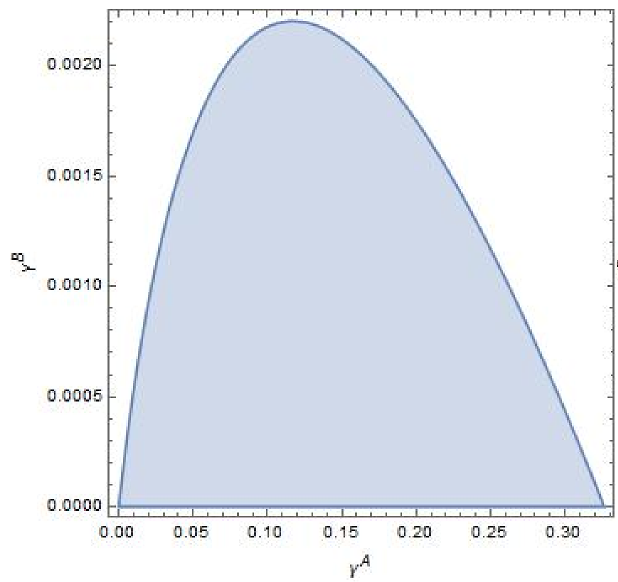
Table 2.1: Possible stable agreements

	Agreement $(1, n)$		Agreement $(2, n)$	
n	$\gamma^A \leq$	$\gamma^B \leq$	$\gamma^A \leq$	$\gamma^B \leq$
3	$\forall$	0.0520	0.0429	$4.85 * 10^{-4}$
4	1.6513	0.0133	0.0223	$1.51 * 10^{-4}$
5	0.8216	0.0070	0.0138	$6.19 * 10^{-5}$
6	0.5461	0.0044	0.0094	$2.99 * 10^{-5}$
7	0.4087	0.0030	0.0068	$1.62 * 10^{-5}$
8	0.3265	0.0022	0.0052	$9.58 * 10^{-6}$
9	0.2718	0.0017	0.0041	$6.01 * 10^{-6}$
10	0.2327	0.0013	0.0033	$3.95 * 10^{-6}$

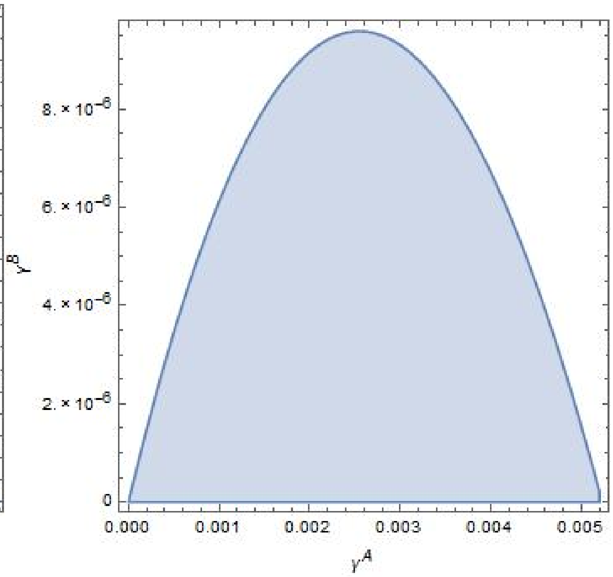
To visualize the results, we consider the two coalitions $(1, 8)$ and $(2, 8)$ and present the corresponding regions, see Figure 2.1, in which the PIS condition (2.29) is satisfied respectively. The X axis shows the parameter γ^A while the Y axis shows the parameter γ^B. The first graph, Figure 2.1a, plots the region where the PIS condition is satisfied so that the agreement $(1, 8)$ is stable (blue area), while the second graph, Figure 2.1b, plots the region where the PIS condition is satisfied so that the agreement $(2, 8)$ is stable (blue area).[6]

Regarding coalition $(1, 8)$, when parameter γ^B takes its maximum value, i.e. $\gamma^B = 0.0022$, the corresponding maximum value that parameter γ^A can take, based

[6]Note that the vertical and horizontal axis' scales are different between the two figures.



(a) Region where $PIS \geq 0$, agreement $(1, 8)$



(b) Region where $PIS \geq 0$, agreement $(2, 8)$

Figure 2.1: Regions where PIS ≥ 0

on condition (2.29), so that the agreement is stable, is $\gamma^A = 0.1218$. Similarly for the coalition $(2, 8)$, when parameter γ^B takes its maximum value, i.e. $\gamma^B = 9.58 * 10^{-6}$, the corresponding maximum value that parameter γ^A can take, so that the agreement is stable, is $\gamma^A = 2.63 * 10^{-3}$. Obviously, when the agreement $(2, 8)$ is stable, the agreement $(1, 8)$ is stable as well since we need stricter constraints for the parameters γ^A and γ^B in order to achieve a stable coalition that includes 2 instead of 1 type A countries. We can present similar graphs for all cases displayed in Table 2.1. The regions where the PIS condition is satisfied take always the same semi-oval form and shrink as we move to larger stable agreements. This is also obvious by comparing the region where the agreement $(1, 8)$ is stable, Figure 2.1a, to the region where the agreement $(2, 8)$ is stable, Figure 2.1b.

Taking the peak points of all curves, like the ones presented in Figure 2.1, for any coalition size, we generate Table 2.2. The table includes the values of the parameters γ^A and γ^B at the peak points.[7] We present also the parameter γ where,

$$\gamma = \frac{\gamma^A}{\gamma^B}, \tag{2.30}$$

and captures the degree of heterogeneity among the two types of countries ($\gamma = \frac{c^A}{c^B}$, given that $b^A = b^B$) for any possible stable coalition.[8]

Table 2.2: Stable agreements for different degrees of heterogeneity

		Agreement $(1, n)$				Agreement $(2, n)$		
n		γ^A	$max\ \gamma^B$	γ		γ^A	$max\ \gamma^B$	γ
3		0.9999	0.0375	26.62		0.01989	$4.85 * 10^{-4}$	40.97
4		0.4557	0.0133	34.13		0.01093	$1.51 * 10^{-4}$	72.32
5		0.2719	0.0070	38.53		0.00679	$6.19 * 10^{-5}$	109.65
6		0.1978	0.0044	44.85		0.00473	$2.99 * 10^{-5}$	157.98
7		0.1468	0.0030	48.53		0.00341	$1.62 * 10^{-5}$	209.84
8		0.1218	0.0022	55.24		0.00263	$9.58 * 10^{-6}$	275.38
9		0.1071	0.0017	63.92		0.00208	$6.01 * 10^{-6}$	347.24
10		0.0920	0.0013	69.65		0.00183	$3.95 * 10^{-6}$	463.15

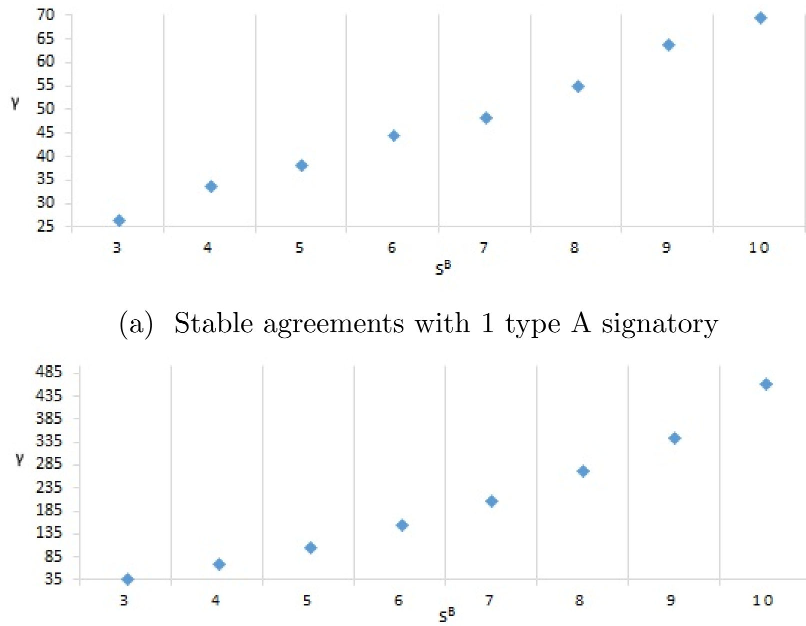
[7] Values are rounded off so that they do not exceed their corresponding maximum points.

[8] Values are rounded to two decimal places.

In all cases, the derived values for the parameters γ^A and γ^B satisfy the following conditions, $0 < \gamma^A < 1$ and $0 < \gamma^B < 1$, except for the agreement $(1, 3)$ where parameter γ^A takes a value higher than 1 when parameter γ^B takes its maximum value. Thus, for this coalition, we restrict $\gamma^A < 1$ and so the corresponding maximum value for γ^B, based on condition (2.29), is $\gamma^B = 0.0375$. The intuition of having $\gamma^j < 1$ is that the slope of the marginal environmental damages (c^j) is smaller than the slope of the marginal benefits (b^j). Therefore, the relative impact of damages to benefits is not very high. In the homogeneous case, the literature has shown that a stable agreement exists, though small, only when the above-mentioned restriction holds (i.e. the impact of damages to benefits is low).

The analysis points out that larger coalitions require stricter constraints for the parameters of the model and a stronger degree of heterogeneity (captured by the parameter γ). Thus, to increase cooperation we have to increase heterogeneity among the two types of countries (higher value for the parameter γ) while decreasing the effect of the global environmental damages on their welfare levels (lower values for the parameters γ^A and γ^B).

Using the data from Table 2.2, we plot in Figure 2.2 the degree of heterogeneity γ (fourth and seventh columns) against the number of type B signatory countries (first column). We display two graphs for the two cases, $(1, n)$ and $(2, n)$ for $n \in \{3, 4, ..., 10\}$. The X axis shows the number of type B signatory countries (i.e. $s^B = n$) and the Y axis shows the parameter γ. In particular, Figure 2.2a depicts the results for the case with 1 type A signatory country and Figure 2.2b depicts the results for the case with 2 type A signatory countries. As indicated, the stronger the degree of heterogeneity, the larger the coalition size.



(a) Stable agreements with 1 type A signatory

(b) Stable agreements with 2 type A signatories

Figure 2.2: Stable agreements for different degrees of heterogeneity

We can demonstrate the fact that larger coalitions are stable only when the degree of heterogeneity increases, by choosing a specific value for the parameter $\gamma^A = 0.0015$ and calculate the required degree of heterogeneity to support different sizes of a stable agreement. Table 2.3 presents the degree of heterogeneity, $\gamma = \gamma^A/\gamma^B$, required, by the PIS condition (2.29), to support stable coalitions consisting of one or two type

A countries (second and third column respectively) and $n \in \{3, 4, ..., 10\}$ type B countries.[9]

Table 2.3: Stable agreements when $\gamma^A = 0.0015$

		Agreement $(1, n)$	Agreement $(2, n)$
		$\gamma^A = 0.0015$	
n		γ	γ
3		4.31	19.14
4		6.85	35.46
5		9.36	58.16
6		11.87	88.97
7		14.36	130.66
8		16.89	187.25
9		19.40	266.40
10		21.93	381.47

The example clearly illustrates that the greater the heterogeneity, the greater the cooperation incentives. For instance, in order to reach the stable agreement with four type B countries and one type A country $(1, 4)$, we need a relatively low level of heterogeneity $\gamma = 6.85$, but in order to have two type A signatories $(2, 4)$, the level of heterogeneity has to be $\gamma = 35.46$. It is worth mentioning that, the degree of heterogeneity required to sustain stable agreements with one type A country increases

[9]Parameter γ is calculated by using the maximum value parameter γ^B takes, given that $\gamma^A = 0.0015$, so that the PIS condition (2.29) is satisfied. Values are rounded to two decimal places.

at a smaller rate as the number of type B signatories increases, relative to the case of agreements with two type A countries.

The above discussion is summarized in following Corollary.

Corollary 7 *A higher degree of heterogeneity is required in order to achieve larger stable agreements. The rate of the required increase in heterogeneity is higher if there are two relative to only one type A signatories.*

The above results extent to any number of countries. The maximum number of type A countries that will join an agreement is two, regardless of their number. Type A signatories, by offering transfers, can attract into the agreement a large number of type B countries, that is increasing with the degree of heterogeneity. In what follows we extent the above results to a larger number of countries such that $n \in \{10, 20, ..., 100\}$.

Following the same process as before, we present in Table 2.4 the values of the parameters γ^A and γ^B at the peak points of the corresponding curves.[10] We include also the parameter γ, i.e. the degree of heterogeneity necessary to support possible stable coalitions such that $s^A \in \{1, 2\}$ and $n \in \{10, 20, ..., 100\}$.[11] Comparing the results presented in Table 2.4 with those displayed in Table 2.2, we observe that the required degree of heterogeneity should be higher in order to induce cooperation of a considerably larger number of type B countries.

[10]Values are rounded off so that they do not exceed their corresponding maximum points.

[11]Values are rounded to two decimal places.

Table 2.4: Larger stable agreements for different degrees of heterogeneity

		Agreement $(1, n)$				Agreement $(2, n)$		
n		γ^A	$max\ \gamma^B$	γ		γ^A	$max\ \gamma^B$	γ
10		0.0920	$1.32 * 10^{-3}$	69.65		$1.83 * 10^{-3}$	$3.95 * 10^{-6}$	463.15
20		0.0378	$2.93 * 10^{-4}$	128.85		$4.16 * 10^{-4}$	$2.53 * 10^{-7}$	$1,647.26$
30		0.0241	$1.25 * 10^{-4}$	191.93		$1.87 * 10^{-4}$	$5.04 * 10^{-8}$	$3,705.38$
40		0.0177	$6.94 * 10^{-5}$	254.92		$1.04 * 10^{-4}$	$1.60 * 10^{-8}$	$6,473.00$
50		0.0141	$4.39 * 10^{-5}$	321.05		$6.73 * 10^{-5}$	$6.58 * 10^{-9}$	$10,221.68$
60		0.0115	$3.03 * 10^{-5}$	379.86		$4.66 * 10^{-5}$	$3.18 * 10^{-9}$	$14,661.68$
70		0.0100	$2.21 * 10^{-5}$	449.21		$3.47 * 10^{-5}$	$1.72 * 10^{-9}$	$20,197.90$
80		0.0087	$1.68 * 10^{-5}$	517.25		$2.67 * 10^{-5}$	$1.01 * 10^{-9}$	$26,519.84$
90		0.0077	$1.33 * 10^{-5}$	575.62		$2.10 * 10^{-5}$	$6.30 * 10^{-10}$	$33,392.52$
100		0.0070	$1.07 * 10^{-5}$	649.54		$1.74 * 10^{-5}$	$4.13 * 10^{-10}$	$42,104.80$

In order to clearly demonstrate the requirement of increasing heterogeneity in order to support larger coalitions, we choose a particular value for the parameter γ^A, that is $\gamma^A = 1.50 * 10^{-5}$, and calculate the value of the parameter γ necessary to support different coalition sizes.[12] Table 2.5 presents the derived results of this exercise.

Summarizing the above discussion, we first find that in order to achieve a larger coalition, a higher degree of heterogeneity is required. Second, the degree of heterogeneity required to sustain stable agreements with one type A country increases at

[12] Parameter γ is calculated by using the maximum value parameter γ^B takes, given that $\gamma^A = 1.50 * 10^{-5}$, so that the PIS condition (2.29) is satisfied. Values are rounded to two decimal places.

Table 2.5: Larger stable agreements when $\gamma^A = 1.50 * 10^{-5}$

		Agreement $(1, n)$	Agreement $(2, n)$
		$\gamma^A = 1.50 * 10^{-5}$	
n		γ	γ
10		21.33	202.95
20		45.52	819.85
30		69.70	1,885.46
40		93.88	3,463.97
50		118.08	5,658.65
60		142.28	8,630.26
70		166.50	12,630.73
80		190.73	18,066.61
90		214.97	25,628.41
100		239.22	36,578.32

a smaller rate as the number of type B signatories increases, relative to the case of agreements with two type A countries.

2.4.2 Transfer rules

We now turn to the design of transfers. Under the optimal transfer rule every coalition member receives at least his free-rider payoff plus a share of the remaining surplus (Eyckmans and Finus, 2004). Therefore, no resources are wasted.

We define the share as $\mu_i^j \geq 0$ such that $\mu = \sum_{i \in S^j, j} \mu_i^j = 1$, where $j \in \{A, B\}$. That is,

$$\mu = \sum_{i=1}^{s^A} \mu_i^A + \sum_{i=1}^{s^B} \mu_i^B = s^A \mu_s^A + s^B \mu_s^B = 1. \tag{2.31}$$

Thus, shares can be different between the two types of countries, however, countries belonging to the same type receive an equal share. This rule is reasonable since countries that benefit from participating in the agreement can induce cooperation without transferring all of their gains to those countries that require compensation for their losses from joining the agreement.

The surplus is defined by the PIS condition in (2.8). Thus, each signatory country receives a share μ_s^j of the surplus,

$$\begin{aligned} PIS(s^A, s^B) &= s^A[\mathcal{W}_s^A(s^A, s^B) - \mathcal{W}_{ns}^A(s^A - 1, s^B)] + \\ &\quad s^B[\mathcal{W}_s^B(s^A, s^B) - \mathcal{W}_{ns}^B(s^A, s^B - 1)]. \end{aligned} \tag{2.32}$$

Every type A coalition member receives final welfare $\mathcal{W}_s^{A^{final}}(s^A, s^B)$,

$$\mathcal{W}_s^{A^{final}}(s^A, s^B) = \mathcal{W}_{ns}^A(s^A - 1, s^B) + \mu_s^A PIS(s^A, s^B), \tag{2.33}$$

and every type B coalition member receives final welfare $\mathcal{W}_s^{B^{final}}(s^A, s^B)$,

$$\mathcal{W}_s^{B^{final}}(s^A, s^B) = \mathcal{W}_{ns}^B(s^A, s^B - 1) + \mu_s^B PIS(s^A, s^B). \tag{2.34}$$

Since type A countries benefit from cooperation, they submit payments, while type B countries receive payments. That is, we have welfare transfers from type A to type B countries, meaning that the first term inside the brackets in condition

(2.32) is positive (internal stability is satisfied for type A countries) while the second term is negative (internal stability is not satisfied for type B countries) for any $s^A \in \{1, 2\}$ and $s^B = n$ with $n \geq 3$. According to the optimal transfer scheme, type A countries should provide each type B signatory its free-rider payoff plus its share of the surplus. Each type A country will also receive its free-rider payoff plus its share of the surplus.

Hence, transfers from type A to type B countries take the following form,

$$T_s(s^A, s^B) = \mathcal{W}_{ns}^B(s^A, s^B - 1) - \mathcal{W}_s^B(s^A, s^B) + \mu_s^B PIS(s^A, s^B). \tag{2.35}$$

At the extreme, type A countries could provide type B countries with just their free-rider payoff, without sharing the surplus. Thus, in this case, $\mu_s^B = 0$, transfers are,

$$T_s(s^A, s^B) = \mathcal{W}_{ns}^B(s^A, s^B - 1) - \mathcal{W}_s^B(s^A, s^B). \tag{2.36}$$

Given the above assumption regarding the transfer rule, the coalition member's welfare after the transfers, is defined in the following Remark.

Remark 8 *After the transfers, the welfare level of type A coalition member is,* $\mathcal{W}_s^{A^{final}}(s^A, s^B) = \mathcal{W}_s^A(s^A, s^B) - \frac{s^B}{s^A} T_s(s^A, s^B)$*, and of type B coalition member is,* $\mathcal{W}_s^{B^{final}}(s^A, s^B) = \mathcal{W}_s^B(s^A, s^B) + T_s(s^A, s^B)$.

2.5 Emissions and welfare levels

The aggregate emissions are given by equation (2.22). Recall that,

$$E = \frac{(a^A n^A + a^B n^B)}{\Psi}. \tag{2.37}$$

Setting $a^A = a^B = a^I$, $b^A = b^B = b^I$, $n^A = n^B = n$ and $s^B = n$, global emissions can be written as,[13]

$$E = \frac{2a^I n}{\Psi}, \tag{2.38}$$

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

Remark 9 *Aggregate emissions decrease in the number of type A signatory countries and in the value of the parameter γ^j, where $j \in \{A, B\}$.*

As expected, when $s^A = 2$ aggregate emissions are lower relative to the case when $s^A = 1$. Moreover, a higher value for the parameter γ^j implies that countries suffer more due to environmental damages and thus tend to emit less.

Proposition 10 *With transfers, large stable agreements emit less. However, the reduction in aggregate emissions achieved by the enlarged agreements is very small relative to the case without transfers.*

Proof. Under the coalition $(s^A = 1, s^B = n)$, global emissions are,[14]

---

[13] The superscript I is used to denote that countries are identical with respect to benefits.

[14] Global emissions are calculated using (2.38) and setting $s^A = 1$.

$$E(s^A = 1, s^B = n) = \frac{2a^I n}{1 + n^2\gamma^B + (\Gamma + \gamma^A)n}. \tag{2.39}$$

Under the coalition ($s^A = 2, s^B = n$), global emissions are,[15]

$$E(s^A = 2, s^B = n) = \frac{2a^I n}{1 + n^2\gamma^B + (\Gamma + \gamma^A)n + 2\gamma^A + \Gamma n}. \tag{2.40}$$

Without transfers, as long as γ^A and γ^B satisfy the appropriate conditions (see Section 1.4.2, Chapter 1), a stable agreement exists such that ($s^A = 2, s^B = 0$). In this case, global emissions are,[16]

$$E(s^A = 2, s^B = 0) = \frac{2a^I n}{1 + \Gamma n + 2\gamma^A}. \tag{2.41}$$

Clearly, for $n \geq 2$,

$$1 + n^2\gamma^B + (\Gamma + \gamma^A)n + 2\gamma^A + \Gamma n > 1 + n^2\gamma^B + (\Gamma + \gamma^A)n > 1 + \Gamma n + 2\gamma^A. \tag{2.42}$$

Therefore,

$$E(s^A = 2, s^B = n) < E(s^A = 1, s^B = n) < E(s^A = 2, s^B = 0). \tag{2.43}$$

■

[15] Global emissions are calculated using (2.38) and setting $s^A = 2$.

[16] Global emissions are calculated using (2.22), restricting heterogeneity in environmental damages and setting $s^A = 2$ and $s^B = 0$.

Table 2.6 presents the global emissions (i.e. E) for the case where $s^A \in \{1, 2\}$ and $n \in \{10, 20, ..., 100\}$.[17] We fix the values for the parameters a^I, b^I, γ^A and γ^B such that $a^I = 1$, $b^I = 25$, $\gamma^A = 1.50 * 10^{-5}$ and $\gamma^B = 4.10 * 10^{-10}$. Given these values for the parameters γ^A and γ^B, all the agreements presented in Table 2.4 are stable. To facilitate comparison the last column of Table 2.6 presents aggregate emissions in the case that no transfers are used and a stable agreement exists such that $(s^A = 2, s^B = 0)$.

Table 2.6: Global emissions

	Transfers		No Transfers
	Agreement $(1, n)$	Agreement $(2, n)$	Agreement $(2, 0)$
n	E	E	E
10	19.994	19.990	19.996
20	39.976	39.963	39.987
30	59.946	59.917	59.971
40	79.904	79.854	79.950
50	99.850	99.772	99.922
60	119.784	119.673	119.889
70	139.706	139.556	139.849
80	159.616	159.421	159.803
90	179.515	179.268	179.752
100	199.401	199.097	199.694

Comparing the second with the third column, it is evident that total emissions are slightly lower with the large agreements $(2, n)$ compared to the agreements $(1, n)$

[17] Values are rounded to three decimal places.

for any corresponding number of n. Comparing the second and the third with the fourth column, it is clear that total emissions are slightly higher in the case without transfers, however, reductions are very small. Thus, even though the presence of transfers increases cooperation, the reduction in aggregate emissions achieved by the enlarged coalitions is very small and so the welfare improvement is also small. Table 2.7 includes the global welfare levels (i.e. W_T) for the cases presented above.[18]

Table 2.7: Global welfare levels

		Transfers			No Transfers
		Agreement $(1, n)$	Agreement $(2, n)$		Agreement $(2, 0)$
n		W_T	W_T		W_T
10		249.250	249.251		249.250
20		494.007	494.011		494.004
30		729.785	729.804		729.769
40		952.112	952.170		952.058
50		$1,156.530$	$1,156.670$		$1,156.390$
60		$1,338.570$	$1,338.860$		$1,338.290$
70		$1,493.810$	$1,494.350$		$1,493.290$
80		$1,617.820$	$1,618.740$		$1,616.930$
90		$1,706.160$	$1,707.630$		$1,704.730$
100		$1,754.440$	$1,756.680$		$1,752.260$

The increase in the coalition size, relative to the case that transfers are not available, comes only from countries belonging to the type with the lower environmental damages (i.e. type B countries), which are drawn into the coalition by the trans-

[18] Values are rounded to three decimal places.

fers offered. The number of coalition members belonging to the type suffering the higher damages (i.e. type A countries) does not increase. Thus, the fact that stable agreements consist of a few countries with high environmental damages and many countries with low environmental damages, confirms the persistent result in the IEAs' literature that large stable coalitions are associated with low gains of cooperation.

2.6 Conclusions

The present paper examines the existence and stability of international environmental coalitions in a two-stage, non-cooperative game among heterogeneous countries while allowing transfers. In particular, we introduce two types of countries differing in their sensitivity to the global pollutant. In order to introduce transfers, the concept of the stability conditions, requiring that none of the coalition's members wish to withdraw from and no country outside the coalition wishes to join the coalition, needs to be modified. We do this by introducing the concept of potential internal stability that allows coalition members to redistribute payoffs among them so that the coalition is internally stable.

We use the usual two-stage emission game where 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. In addition, in the first stage those countries that decide to join the agreement agree also to share the gains from cooperation. We apply the following optimal transfer rule: type A countries give every type B country, member of the coalition, his free-rider payoff and they share the remaining gains among either all members or themselves.

Our results show that allowing for transfers can increase cooperation among heterogeneous countries. Although the increase in the coalition size can be considerable, the coalition's expansion is based only on countries of type B drawn into the coalition by the incentive of the transfers offered by countries of type A which suffer the higher environmental damages. Type A countries' free-riding incentives are strong and thus the coalition does not expand by including more of them. Since the coalition contains more type B countries, that they do not have strong incentives to decrease emissions, the reduction in aggregate emissions due to the enlargement of the coalition is small, leading to dismal improvement in welfare.

Consequently, based on our analysis, using simulations, we can conclude that a stable with transfers agreement can have either one or two type A countries and any number n of type B countries. The level of cooperation that can be achieved using transfers increases with the degree of heterogeneity, meaning that the higher the heterogeneity in environmental damages, the higher the level of cooperation. Furthermore, with transfers large stable coalitions can perform only slightly better in terms of reductions in emissions.

2.7 Appendix

In what follows we present the proofs of Remarks 8 and 9.

Proof of Remark 8. Under the optimal transfer rule, every coalition member receives at least his free-rider payoff plus a share of the remaining surplus. Based on our analysis, type A countries should give each type B country - member of the coalition - his free-rider payoff. They will each receive also their free-rider payoffs and then share the remaining gains among all members.

Recall that,

$$\begin{aligned} PIS(s^A, s^B) &= s^A[\mathcal{W}_s^A(s^A, s^B) - \mathcal{W}_{ns}^A(s^A - 1, s^B)] + \\ &\quad s^B[\mathcal{W}_s^B(s^A, s^B) - \mathcal{W}_{ns}^B(s^A, s^B - 1)] \geq 0. \end{aligned} \tag{2.44}$$

Transfers can take the following form,

$$T_s(s^A, s^B) = \mathcal{W}_{ns}^B(s^A, s^B - 1) - \mathcal{W}_s^B(s^A, s^B) + \mu_s^B PIS(s^A, s^B). \tag{2.45}$$

The total transfers that should be paid to type B coalition members are,

$$\begin{aligned} T_s^{total}(s^A, s^B) &= s^B \left[\mathcal{W}_{ns}^B(s^A, s^B - 1) - \mathcal{W}_s^B(s^A, s^B) + \mu_s^B PIS(s^A, s^B)\right] \\ &= -s^B \left[\mathcal{W}_s^B(s^A, s^B) - \mathcal{W}_{ns}^B(s^A, s^B - 1)\right] + (1 - s^A \mu_s^A) PIS(s^A, s^B) \\ &= s^A \left[\mathcal{W}_s^A(s^A, s^B) - \mathcal{W}_{ns}^A(s^A - 1, s^B)\right] - s^A \mu_s^A PIS(s^A, s^B). \end{aligned} \tag{2.46}$$

Each type A country should pay,

$$\frac{T_s^{total}(s^A, s^B)}{s^A} = \mathcal{W}_s^A(s^A, s^B) - \mathcal{W}_{ns}^A(s^A - 1, s^B) - \mu_s^A PIS(s^A, s^B). \tag{2.47}$$

Therefore, the final welfare for each type A country is,

$$\begin{aligned} \mathcal{W}_s^{A^{final}}(s^A, s^B) &= \mathcal{W}_s^A(s^A, s^B) - \frac{T_s^{total}(s^A, s^B)}{s^A} \\ &= \mathcal{W}_{ns}^A(s^A - 1, s^B) + \mu_s^A PIS(s^A, s^B). \end{aligned} \tag{2.48}$$

Moreover, the final welfare for each type B country is,

$$\begin{aligned}
\mathcal{W}_s^{B^{final}}(s^A, s^B) &= \mathcal{W}_s^B(s^A, s^B) + T_s(s^A, s^B) \\
&= \mathcal{W}_s^B(s^A, s^B) + \mathcal{W}_{ns}^B(s^A, s^B - 1) - \mathcal{W}_s^B(s^A, s^B) + \mu_s^B PIS(s^A, s^B) \\
&= \mathcal{W}_{ns}^B(s^A, s^B - 1) + \mu_s^B PIS(s^A, s^B). \qquad (2.49)
\end{aligned}$$

In the extreme case where type A countries give every type B country only his free-rider payoff without any share of the remaining surplus, parameters μ_s^B should be equal to zero, i.e. $\mu_s^B = 0$. Thus, transfers can be simplified as follows,

$$T_s(s^A, s^B) = \mathcal{W}_{ns}^B(s^A, s^B - 1) - \mathcal{W}_s^B(s^A, s^B). \qquad (2.50)$$

Since gains are distributed only among type A countries, $\mu_s^A = \frac{1}{s^A}$. That is, $\mu = \sum_{i \in S^A} \mu_i^A = s^A \mu_s^A = 1$. Hence, every type A coalition member receives final welfare,

$$\begin{aligned}
\mathcal{W}_s^{A^{final}}(s^A, s^B) &= \mathcal{W}_{ns}^A(s^A - 1, s^B) + \frac{1}{s^A} PIS(s^A, s^B) \\
&= \mathcal{W}_s^A(s^A, s^B) + \frac{s^B}{s^A}\left[\mathcal{W}_s^B(s^A, s^B) - \mathcal{W}_{ns}^B(s^A, s^B - 1)\right] \\
&= \mathcal{W}_s^A(s^A, s^B) - \frac{s^B}{s^A} T_s(s^A, s^B). \qquad (2.51)
\end{aligned}$$

Every type B coalition member receives final welfare,

$$\begin{aligned}
\mathcal{W}_s^{B^{final}}(s^A, s^B) &= \mathcal{W}_{ns}^B(s^A, s^B - 1) \\
&= \mathcal{W}_s^B(s^A, s^B) + T_s(s^A, s^B). \qquad (2.52)
\end{aligned}$$

■

Proof of Remark 9. The aggregate emissions can be written as,

$$E = \frac{2a^I n}{\Psi}, \tag{2.53}$$

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

The derivative of the aggregate emissions with respect to the number of type A signatory countries, i.e. s^A, is negative meaning that global emissions decrease in the number of type A signatory countries.

$$D_{s^A} = -2a^I n \frac{\gamma^A(2s^A - 1) + \Gamma n}{\Psi^2}. \tag{2.54}$$

The derivative of the aggregate emissions with respect to the parameter γ^A is negative. When parameter γ^A increases, type A countries suffer more due to environmental pollution and thus tend to emit less.

$$D_{\gamma^A} = -2a^I n \frac{n + s^A(s^A + n - 1)}{\Psi^2}. \tag{2.55}$$

The derivative of the aggregate emissions with respect to the parameter γ^B is negative. When parameter γ^B increases, type B countries suffer more due to environmental pollution and thus tend to emit less.

$$D_{\gamma^B} = -2a^I n \frac{n(s^A + n)}{\Psi^2}. \tag{2.56}$$

■

CHAPTER 3

INTERNATIONAL ENVIRONMENTAL AGREEMENTS AND TRADING BLOCKS - CAN ISSUE LINKAGE ENHANCE COOPERATION?

3.1 Introduction

The present chapter examines IEAs in an economy with trade.[1] We extent the basic (non-cooperative) model of the IEAs' literature with quadratic cost and benefit functions and simultaneous decisions by letting countries choose emission taxes and import tariffs as their policy instruments to manage climate change and control trade. We consider the formation of a Global Agreement (GA) where countries (named signatories) that form an environmental agreement form a free trade agreement as well. Nations that remain outside of the agreement (named non-signatories) suffer trade costs. The advantage for the signatories is the tariff-free access to other signatories' markets while at the same time they bare the burden of reducing emissions more. In contrast, the disadvantage for the non-signatories is that they have to pay tariffs on their imports to any other country while free-riding on environmental efforts.

In this context, we define the equilibrium of a three-stage emission game. In the first stage, each country decides whether or not to join the agreement. In the second

[1] This chapter is a joint work with my supervisor Dr. Effrosyni Diamantoudi (Department of Economics, Concordia University).

stage, countries choose simultaneously - cooperatively or non-cooperatively - tariff and tax levels. In the third stage, each firm, taking the policies set by the countries and the output decisions of the other firms as given, maximizes its profits. To obtain the subgame perfect Nash equilibrium, the model is solved by backward induction and stability is determined by applying the internal and external stability conditions (D'Aspremont et al., 1983) as well as the admissibility condition.

The main objective of the present study is to investigate the effect of trade on the stability and effectiveness of environmental agreements, when all non-cooperative countries and coalition members choose both their terms of trade and climate policy instruments to deal with the environmental pollution. In particular, we are interested in examining whether the presence of trade can enhance cooperation and improve environmental performance as well as welfare relative to the basic model of the IEAs' literature.

Our analysis is mostly related to Eichner and Pethig (2015) in the sense that we study IEAs in a model with symmetric countries, international trade, and emission tax policy, however, they model a free trade world economy. To our knowledge, there are no relevant studies that examine the formation of GAs in a framework similar to ours. We believe that the existence of the two instruments (i.e. tariffs and taxes), even though adds complexity to the analysis, captures better the real-world situation since trade and environmental problems affect each other. Under the formation of a GA, there are two important effects that have to be taken into consideration. In an IEA, the coalition formation creates positive externalities on non-participants. On the other hand, in trade agreements, the coalition formation creates negative externalities on non-participants (Yi, 1997). The interaction between these two effects is essential to determine the stability and effectiveness of an agreement.

Results are optimistic. Our findings illustrate the importance of trade and environmental policies working together to improve cooperation in effective agreements. Contrary to the IEA model, stable agreements are larger and more efficient in reducing aggregate emissions and improving welfare. Moreover, the analysis shows that the size of a stable agreement increases in the number of countries affected by the externalities. This result appears to contradict the main conclusion reached in the IEAs' literature stating that in a Cournot-IEA the maximum level of cooperation consists of two members, independently of the number of countries. Clearly, if world markets do not exist (autarky), the model coincides with the basic model of the IEAs' literature.

The rest of the chapter is structured as follows. Section 3.2 describes the model. Section 3.3 presents the benchmark case. Section 3.4 examines the formation of a GA. Section 3.5 presents the stability conditions. Section 3.6 analyses numerically the effect of trade on the stability and effectiveness of environmental agreements. Section 3.7 concludes.

3.2 The model

We consider an open economy where countries trade with each other. We assume that there are n identical countries, $N = \{1, 2, 3, ..., n\}$. The representative consumer in each country $i \in N$ has a utility function of the form,

$$U_i(e_i^c; K_i) = b\left(ae_i^c - \frac{1}{2}(e_i^c)^2\right) + K_i, \tag{3.1}$$

where e_i^c is country i's total consumption of the nonnumeraire good, K_i denotes the numeraire good, and a and b are positive parameters, i.e. $a > 0$ and $b > 0$. The total consumption e_i^c is given by,

$$e_i^c = \sum_{j=1}^{n} e_{ij} = e_i^d + \sum_{j \neq i} e_{ij}^I, \tag{3.2}$$

where e_i^d is country i's consumption of the domestic product and e_{ij}^I indicates the quantity country i imports from country $j \neq i$ (i.e. quantity sold from country j to country i).

The quasilinear utility function, given by equation (3.1), is sufficient to derive country i's inverse demand function for the nonnumeraire good, that is,

$$p_i(e_i^c) = b(a - e_i^c). \tag{3.3}$$

The numeraire good is produced under perfect competition with constant returns to scale and the nonnumeraire good is produced, in each country, by a single profit maximizing firm. For simplicity, the marginal cost of production is assumed to be constant and equal to zero. We consider that there is no pollution associated with the numeraire good, while each unit of the nonnumeraire good produced generates one unit of pollution emission.

Country i charges a non-negative tariff at the same rate of τ_i per unit of imports from any country j, where $j \neq i$. Then country j's effective marginal cost of exporting to country i is τ_i. Similarly, country i's effective marginal cost of exporting to country j is τ_j. We consider trade only in the good that generates emissions (i.e. nonnumeraire good).

A by-product of production in this model is pollution. Pollution is perfectly transboundary and thus affects widely all countries. Country i's production and as a consequence polluting activity (recall the one-by-one relationship between production of consumption good and pollution) is given by,

$$e_i^p = e_i^d + \sum_{j \neq i} e_{ij}^X. \tag{3.4}$$

That is, production in country i is the sum of what the country produces and consumes domestically (i.e. e_i^d) and what the country produces domestically and exports (i.e. e_{ij}^X).

The damage from pollution is monotonically increasing and convex in the global emissions, $E = \sum_{i=1}^{n} e_i^p$. In particular, the damage function is given by,

$$D_i(E) = \frac{1}{2} c E^2, \tag{3.5}$$

where $c > 0$ is the pollution damage parameter, as well as $D_i(0) = 0$, $D_i' \geq 0$ and $D_i'' > 0$. A convex environmental damage function implies that damages from emissions increase at an increasing rate and so gradually reduce ecosystem services. In other works, more emissions cause more harm on nature.

The environmental policy in country i is a carbon tax imposed per unit of emission by the domestic firm due to its production. Given our assumption that each unit of the polluting good produced generates one unit of pollution emission, a tax per unit of emission is equivalent to a tax per unit of the polluting good. The pollution (or emission) tax set in country i is denoted by t_i.

We model the process of countries' decision as a non-cooperative three-stage emission game and start by solving the third stage where, taking countries' decisions as given, firms compete a la Cournot in the product markets. In the second stage, countries choose simultaneously - cooperatively or non-cooperatively - tariff and tax levels and in the first stage, each country decides whether or not to join the agreement.

Firm's problem

Each firm maximizes profits taking the policies set by the countries and the output decisions of the other firms as given. The total profits for the firm i located in country i consist of the profits of sales in the domestic market i plus the profits of sales (i.e. profits of output exported) in country j, minus the pollution tax imposed on emissions. Thus,

$$\Pi_i = \Pi_i^d + \left(\sum_{j \neq i} \Pi_{ij}^X \right) - t_i e_i^p. \tag{3.6}$$

Firm i maximizes profits, given by equation (3.6), by choosing quantity sold in country i, i.e. e_i^d, and quantity sold in country j, i.e. e_{ij}^X, for all $j \neq i \in N$. Given that firm i has a zero marginal cost of producing the homogeneous good while its effective marginal cost of exporting to country j is τ_j, its profits can be written as,

$$\Pi_i = p_i(e_i^c) e_i^d + \sum_{j \neq i} (p_j(e_j^c) - \tau_j) e_{ij}^X - t_i e_i^p, \tag{3.7}$$

where $e_j^c = \sum_{i=1}^n e_{ji}$ is country j's total consumption of the nonnumeraire good.

The maximization problem is,

$$\max_{e_i^d,\, e_{ij}^X} \Pi_i = p_i(e_i^c)e_i^d + \sum_{j\neq i}(p_j(e_j^c) - \tau_j)e_{ij}^X - t_i e_i^p, \forall j \neq i \in N. \tag{3.8}$$

Country's problem

We assume that firms' profits and tariff and tax revenues are rebated back to the consumers. So that, country i's welfare, denoted by W_i, consists of the domestic consumer surplus CS_i, the domestic firm's profits Π_i (net of all taxes), the tariff revenues TR_i, the tax revenues ER_i and the environmental damages due to the aggregate pollution level $D_i(E)$.[2] That is,

$$W_i = CS_i + \Pi_i + TR_i + ER_i - D_i(E). \tag{3.9}$$

Thus, country i's total welfare can be written as,

$$W_i = \frac{1}{2}b\,(e_i^c)^2 + p_i(e_i^c)e_i^d + \sum_{j\neq i}(p_j(e_j^c) - \tau_j)e_{ij}^X + \tau_i \sum_{j\neq i} e_{ij}^I - \frac{1}{2}cE^2, \tag{3.10}$$

where the quantity e_{ij}^X indicates country i's exports to country j while the quantity e_{ij}^I indicates country's i imports from country j.

Country i maximizes welfare given by (3.10) by choosing tariff level τ_i and tax level t_i. Thus, its maximization problem is,

[2] The consumer surplus CS_i is calculated by taking, $U_i(e_i^c; K_i) - p_i(e_i^c)e_i^c$.

$$\max_{\tau_i, t_i} W_i = \frac{1}{2} b \left(e_i^c\right)^2 + p_i(e_i^c) e_i^d + \sum_{j \neq i} (p_j(e_j^c) - \tau_j) e_{ij}^X + \tau_i \sum_{j \neq i} e_{ij}^I - \frac{1}{2} c E^2. \tag{3.11}$$

3.3 The benchmark case

The non-cooperative outcome arises when each country $i \in N$ chooses its tariff and tax levels taking as given the tariff and tax levels from all the other countries, playing Nash equilibrium.

From the firms' maximization problem (3.8), the first order condition with respect to the domestic quantity e_i^d gives the following expression,

$$ab - b \sum_{j=1}^{n} e_{ij} - b e_i^d - t_i = 0. \tag{3.12}$$

Moreover, the first order condition with respect to the quantity exported e_{ij}^X gives the following expression,[3]

$$ab - b \sum_{i=1}^{n} e_{ji} - b e_{ij}^X - t_i - \tau_j = 0. \tag{3.13}$$

Using the first order conditions (3.12) and (3.13), we derive country i's reaction functions for the domestic product and the quantity exported in the benchmark case. Thus,

---

[3]The quantity sold (exported) from country i to country j indicates also the quantity country j imports from country i, that is $e_{ij}^X = e_{ji}^I$.

$$e_i^d = \frac{ab - t_i - b(n-1)e_{ij}^I}{2b}, \tag{3.14}$$

$$e_{ij}^X = \frac{ab - t_i - \tau_j - be_j^d}{nb}. \tag{3.15}$$

Due to symmetry, the reaction function for the quantity exported from country j to country i (i.e. e_{ji}^X), that is the quantity imported to country i, is given by,

$$e_{ij}^I = \frac{ab - t_j - \tau_i - be_i^d}{nb}. \tag{3.16}$$

Using the reaction functions, we derive the following expressions,

$$e_i^d = \frac{ab - nt_i + (n-1)t_j + (n-1)\tau_i}{(n+1)b}, \tag{3.17}$$

$$e_{ij}^X = \frac{ab + t_j - 2t_i - 2\tau_j}{(n+1)b}, \tag{3.18}$$

$$e_{ij}^I = \frac{ab + t_i - 2t_j - 2\tau_i}{(n+1)b}, \tag{3.19}$$

$$e_i^c = \frac{abn - t_i - (n-1)t_j - (n-1)\tau_i}{(n+1)b}. \tag{3.20}$$

Country i's maximization problem (3.11) can be written as,

$$\max_{\tau_i, t_i} W_i = \frac{1}{2} b \left(e_i^c\right)^2 + p_i(e_i^c) e_i^d + (n-1)(p_j(e_j^c) - \tau_j) e_{ij}^X + (n-1)\tau_i e_{ij}^I - \frac{1}{2} c E^2. \quad (3.21)$$

Before we proceed to the solutions, we define parameter γ as the ratio between environmental damages and benefits due to emissions. Therefore,

$$\gamma = \frac{c}{b}. \quad (3.22)$$

The first order conditions for the welfare maximization yield two reaction functions corresponding to the equilibrium τ_i and t_i. Since countries are identical, the tariff and pollution tax will be identical for all countries. Hence, imposing $\tau_i = \tau_j$ and $t_i = t_j$ in country i's reaction functions we solve for the Nash equilibrium tariff and tax. The reaction functions (after imposing $\tau_i = \tau_j$ and $t_i = t_j$) are presented in Appendix A.

Therefore, we have,

$$\tau_i = ab \frac{(2(2n-1)\gamma + 1)n^2 + 3n - 2}{2(n+1)(3n + (2n-1)n\gamma - 2)}, \quad (3.23)$$

$$t_i = ab \frac{(4(2n-1)\gamma + n - 6)n + 1}{2(n+1)(3n + (2n-1)n\gamma - 2)}. \quad (3.24)$$

The domestic quantity is given by,

$$e_i^d = a\frac{2((2n-3)n+1)n\gamma + (n+6)n - 3}{2(n+1)(3n+(2n-1)n\gamma - 2)}. \tag{3.25}$$

The quantity imported from country i is equal to its quantity exported.[4] That is,

$$e_{ij}^I = e_{ij}^X = a\frac{3n - 2n(2n-1)\gamma - 1}{2(n+1)(3n+(2n-1)n\gamma - 2)}. \tag{3.26}$$

The total quantity consumed in country i is equal to its total quantity produced. That is,

$$e_i^c = e_i^p = a\frac{2n-1}{3n+(2n-1)n\gamma - 2}. \tag{3.27}$$

The aggregate consumption level (which is equal to the aggregate production level) is,

$$E_{nc} = an\frac{2n-1}{3n+(2n-1)n\gamma - 2}. \tag{3.28}$$

Given our assumption that each unit of the nonnumeraire good produced generates one unit of pollution emission, equation (3.28) represents the aggregate emissions as well.

Country's i welfare is given by,

$$\mathcal{W}_i = a^2 b\frac{(2n-1)(4n-((2n-5)n+2)n\gamma - 3)}{2(3n+(2n-1)n\gamma - 2)^2}. \tag{3.29}$$

---

[4] Imports are non-negative for all $n > 0$ and $\gamma \leq \frac{3n-1}{2n(2n-1)}$.

The world market-clearing condition, which requires that global production of the good to be equal to its global consumption, is satisfied.

3.4 Coalition formation

We consider that a set of countries signs an GA aiming at controlling emissions and trade. We call those countries signatories, while the non-participants of the agreement are called non-signatories. Signatories trade freely among themselves, while non-signatories are penalized by a tariff on their imports to the members of the agreement. Moreover, non-signatories pay a tariff when they export to other non-signatories. Additionally, signatories choose a common tax level that internalizes the full environmental cost of all coalition members while non-signatories choose individually their emission tax.

In particular, we assume that a set of countries $S \subset N$ signs a GA and the remaining $N \backslash S$ do not. The countries that form an agreement of size $s = |S|$, act cooperatively maximizing the joint welfare, while the countries that decide not to participate act non-cooperatively maximizing their own welfare. Thus, there are s signatory countries and $(n - s)$ non-signatory countries. Taking advantage of the symmetry assumption, we treat all signatory countries equally within the coalition.

3.4.1 Output levels

A signatory country's total consumption of the nonnumeraire good is given by,

$$e_s^c = e_s^d + (s-1)e_{ss}^I + (n-s)e_{sns}^I, \tag{3.30}$$

where e_s^d is signatory's domestic product, e_{ss}^I indicates quantity imported from a signatory country, and e_{sns}^I indicates quantity imported from a non-signatory country.

A non-signatory country's total consumption of the nonnumeraire good is given by,

$$e_{ns}^c = e_{ns}^d + (n - s - 1)e_{nsns}^I + se_{nss}^I, \tag{3.31}$$

where e_{ns}^d is non-signatory's domestic product, e_{nsns}^I indicates quantity imported from a non-signatory country, and e_{nss}^I indicates quantity imported from a signatory country.

Using the first order conditions (3.12) and (3.13) from the firm's maximization problem, we derive the domestic product and the quantities imported for each signatory and non-signatory country respectively.

For a signatory country, these quantities take the following forms,[5]

$$e_s^d = e_{ss}^I = \frac{ab - (n - s + 1)t_s + (n - s)t_{ns} + (n - s)\tau_s}{(n + 1)b}, \tag{3.32}$$

$$e_{sns}^I = \frac{ab + st_s - (s + 1)t_{ns} - (s + 1)\tau_s}{(n + 1)b}. \tag{3.33}$$

We restrict the parameter values so that imports are non-negative. That is, $ab \geq (s + 1)(t_{ns} + \tau_s) - st_s$. If a signatory country raises its tariff on imports from a non-signatory country, i.e. τ_s, then production (as well as consumption) of the

[5] The quantity a signatory country imports from a non-signatory country indicates also the quantity a non-signatory country exports to a signatory country, that is $e_{sns}^I = e_{nss}^X$. Also, signatories exchange an equal quantity among themselves, that is $e_{ss}^I = e_{ss}^X$.

domestic good increases, imports from another signatory country increase as well, while imports from a non-signatory country decrease. Similar effects occur when a non-signatory country increases its tax per unit of emission, i.e. t_{ns}. However, an increase in a signatory country's tax per unit of emission, i.e. t_s, will cause a decrease in domestic production and consumption, a decrease in imports from a signatory country and an increase in imports from a non-signatory country.

To derive a signatory country's total consumption of the nonnumeraire good, i.e. e_s^c, we use equations (3.32) and (3.33). Therefore,

$$e_s^c = \frac{abn - st_s - (n-s)t_{ns} - (n-s)\tau_s}{(n+1)b}. \tag{3.34}$$

If a signatory country raises its tariff on imports from a non-signatory country, i.e. τ_s, then its total consumption falls. An increase in either a signatory country's tax per unit of emission, i.e. t_s, or a non-signatory country's tax per unit of emission, i.e. t_{ns}, will cause a decrease to its total consumption as well.

For a non-signatory, these quantities take the following forms,[6]

$$e_{ns}^d = \frac{ab - (s+1)t_{ns} + st_s + (n-1)\tau_{ns}}{(n+1)b}, \tag{3.35}$$

$$e_{nsns}^I = \frac{ab - (s+1)t_{ns} + st_s - 2\tau_{ns}}{(n+1)b}, \tag{3.36}$$

[6]The quantity a non-signatory country imports from a signatory country indicates also the quantity a signatory country exports to a non-signatory country, that is $e_{nss}^I = e_{sns}^X$. Also, non-signatories exchange an equal quantity among themselves, that is $e_{nsns}^I = e_{nsns}^X$.

$$e^I_{nss} = \frac{ab + (n-s)t_{ns} - (n-s+1)t_s - 2\tau_{ns}}{(n+1)b}. \tag{3.37}$$

We restrict the parameter values so that imports are non-negative. That is, $ab \geq (s+1)t_{ns} + 2\tau_{ns} - st_s$ and $ab \geq (n-s+1)t_s + 2\tau_{ns} - (n-s)t_{ns}$. If a non-signatory country raises its tariff on imports from another country, i.e. τ_{ns}, then production (as well as consumption) of the domestic good increases, while imports decrease. An increase in a non-signatory country's tax per unit of emission, i.e. t_{ns}, will cause a reduction in domestic quantity and imports from another non-signatory country but an increase in imports from a signatory country. The inverse effect occurs if a signatory country's tax per unit of emission, i.e. t_s, increases. That is, domestic quantity and imports from another non-signatory country increase but imports from a signatory country decrease.

Using equations (3.35), (3.36) and (3.37), we derive the total consumption of a non-signatory country. That is,

$$e^c_{ns} = \frac{abn - (n-s)t_{ns} - st_s - (n-1)\tau_{ns}}{(n+1)b}. \tag{3.38}$$

If a non-signatory country raises its tariff on imports, i.e. τ_{ns}, then its total consumption falls. An increase in either non-signatory country's tax per unit of emission, i.e. t_{ns}, or a signatory country's tax per unit of emission, i.e. t_s, will cause a decrease to its total consumption as well.

A signatory country's net imports are,

$$(n-s)(e^I_{sns} - e^X_{sns}) = -(n-s)\frac{(n+1)t_{ns} - (n+1)t_s - 2\tau_{ns} + (s+1)\tau_s}{(n+1)b}. \tag{3.39}$$

A non-signatory country's net imports are,

$$s(e^I_{nss} - e^X_{nss}) = s\frac{(n+1)t_{ns} - (n+1)t_s - 2\tau_{ns} + (s+1)\tau_s}{(n+1)b}. \tag{3.40}$$

Thus, global net imports sum to zero clearing the markets. That is,

$$s(n-s)(e^I_{sns} - e^X_{sns}) + (n-s)s(e^I_{nss} - e^X_{nss}) = 0. \tag{3.41}$$

Using equations (3.34) and (3.38), we derive the aggregate consumption level which is equal to the aggregate production level (due to the world market-clearing condition). That is,

$$E = s\frac{abn - st_s - (n-s)t_{ns} - (n-s)\tau_s}{(n+1)b} + (n-s)\frac{abn - (n-s)t_{ns} - st_s - (n-1)\tau_{ns}}{(n+1)b}. \tag{3.42}$$

Rearranging,

$$E = \frac{abn^2 - nst_s - n(n-s)t_{ns} - s(n-s)\tau_s - (n-1)(n-s)\tau_{ns}}{(n+1)b}. \tag{3.43}$$

Given our assumption that each unit of the nonnumeraire good produced generates one unit of pollution emission, equation (3.43) represents the aggregate emissions as well. We observe that aggregate emissions decrease when the tariff and tax levels set by either signatories or non-signatories or both parties increase.

3.4.2 Tariff and tax levels

Under the formation of a GA of size $s = |S|$, signatories abolish tariffs among themselves and jointly choose their external tariff (i.e. tariff to the non-signatories) and tax level to maximize the aggregate welfare of the members. On the other hand, non-signatories choose their tariff and emission tax to maximize their own welfare.

Signatories maximize the aggregate coalition welfare $\sum_{i \in S} W_i = sW_s$ with respect to τ_s and t_s. Their maximization problem is,

$$\max_{\tau_s, t_s} sW_s = s \left(\begin{array}{c} \frac{1}{2}b\left(e_s^c\right)^2 + sp_s(e_s^c)e_s^d + (n-s)(p_{ns}(e_{ns}^c) - \tau_{ns})e_{sns}^X + \\ (n-s)\tau_s e_{sns}^I - \frac{1}{2}cE^2 \end{array} \right). \tag{3.44}$$

Note that a signatory's profits from exporting to other signatories are equal to its domestic profits since $e_s^d = e_{ss}^X$ as per equation (3.32). Moreover, it receives profits from exporting to the non-signatories after taking into account the export related costs. There are also tariff revenues per unit of imports from the non-signatories.

The first order conditions for the welfare maximization yield two reaction functions. The signatories' reaction function for the equilibrium tariff $\tau_s(t_s, \tau_{ns}, t_{ns})$ which is a function of signatories' tax t_s, non-signatories' tariff τ_{ns}, non-signatories' tax t_{ns}, and the other parameters in the model. The signatories' reaction function for the equilibrium tax $t_s(\tau_s, \tau_{ns}, t_{ns})$ which is a function of signatories' tariff τ_s, non-signatories' tariff τ_{ns}, non-signatories' tax t_{ns}, and the other parameters in the model. The corresponding second order conditions for the welfare maximization problem are satisfied (see Appendix B).

Non-signatories maximize their own welfare W_{ns} with respect to τ_{ns} and t_{ns}. Their maximization problem is,

$$\max_{\tau_{ns},t_{ns}} W_{ns} = \left(\begin{array}{c} \frac{1}{2}b\left(e_{ns}^{c}\right)^{2} + p_{ns}(e_{ns}^{c})e_{ns}^{d} + (n-s-1)p_{ns}(e_{ns}^{c})e_{nsns}^{X} + \\ s(p_{s}(e_{s}^{c}) - \tau_{s})e_{nss}^{X} + s\tau_{ns}e_{nss}^{I} - \frac{1}{2}cE^{2} \end{array} \right). \qquad (3.45)$$

A non-signatory's profits from exporting to other non-signatories are different from its domestic profits since $e_{ns}^{d} \neq e_{nsns}^{X}$ as per equations (3.35) and (3.36). Notice that the costs related to those exports are equal to the tariff revenues from the other non-signatories' imports since they all exchange an equal quantity among themselves. Additionally, it receives profits from exporting to the signatories after taking into account the export costs. There are also tariff revenues per unit of imports from the signatories.

The first order conditions for the welfare maximization yield two reaction functions. The non-signatories' reaction function for the equilibrium tariff $\tau_{ns}(t_{ns}, \tau_{s}, t_{s})$ which is a function of non-signatories' tax t_{ns}, signatories' tariff τ_{s}, signatories' tax t_{s}, and the other parameters in the model. The non-signatories' reaction function for the equilibrium tax $t_{ns}(\tau_{ns}, \tau_{s}, t_{s})$ which is a function of non-signatories' tariff τ_{ns}, signatories' tariff τ_{s}, signatories' tax t_{s}, and the other parameters in the model. The corresponding second order conditions for the welfare maximization problem are satisfied (see Appendix B).

The reaction functions for the tariff, τ_s and τ_{ns}, and tax, t_s and t_{ns}, levels are presented in Appendix C. Moreover, the equilibrium levels of tariffs and taxes are presented in Appendix D due to the length of their expressions.

3.4.3 Full cooperation case

In the full cooperation case, countries abolish tariffs and jointly choose their tax level to maximize the aggregate welfare. Under the full cooperation assumption, tariffs are eliminated from our analysis and all countries choose the tax t_c that maximizes aggregate welfare. In this case, the model is simplified to the basic full cooperation model (socially optimal outcome) of the IEAs' literature. As a result the tax level imposed by a country is given by,

$$t_c = \frac{ab(n^3\gamma - 1)}{n(1 + n^2\gamma)}. \tag{3.46}$$

The quantities produced and traded are all equal and thus total quantity consumed in each country is equal to its total quantity produced. That is,

$$e_c^c = e_c^p = n\left(\frac{a}{n + \gamma n^3}\right). \tag{3.47}$$

The aggregate emissions are given by,

$$E_c = \frac{an}{1 + \gamma n^2}. \tag{3.48}$$

Each country receives welfare given by,

$$\mathcal{W}_c = \frac{a^2 b}{2(1+\gamma n^2)}. \tag{3.49}$$

The derived solutions for the total emissions and welfare are equivalent to the solutions presented in Rubio and Casino (2001) where they calculate the full cooperative level of emissions and net benefits of each country in the IEA model.[7]

3.5 The stability of an agreement

In the IEAs' literature, the existence and stability of an environmental agreement is determined using the notions of internal and external stability as was originally developed by D'Aspremont et. al (1983) and extended to IEAs by Carraro and Siniscalco (1993) and Barrett (1994). Internal stability implies that no coalition member has an incentive to leave the coalition, while external stability implies that no country outside the coalition has an incentive to join the coalition, assuming that the remaining players in the game do not revise their membership decision. We denote the size of a stable agreement by s^*.

Formally, the internal and external stability conditions take the following forms:

internal stability condition,

$$\mathcal{W}_s(s^*) \geq \mathcal{W}_{ns}(s^*-1), \tag{3.50}$$

[7] In our model, a representative consumer in country i has a utility function of the form $U_i(e_i^c; K_i) = b\left(ae_i^c - \frac{1}{2}\left(e_i^c\right)^2\right) + K_i$ while Rubio and Casino (2001) assume that the quadratic benefit function for each country takes the form, $B_i(q_i) = aq_i - \frac{b}{2}q_i^2$, where q_i denotes emissions by country i, $a > 0$ and $b > 0$. It is trivial to derive the equivalence between the parameters.

external stability condition,

$$\mathcal{W}_{ns}(s^*) \geq \mathcal{W}_s(s^* + 1), \tag{3.51}$$

where $\mathcal{W}_s$ is the welfare of a signatory country and $\mathcal{W}_{ns}$ is the welfare of a non-signatory country.

The present model examines the formation of IEAs in an economy with trade. Therefore, we have to take into consideration two effects. The first effect, referred to as environmental effect, is related to the pure public good provision problem. Since environment is a global public good, countries have strong free-riding incentives especially when compliance with an IEA is costly to them. That is, the coalition formation generates positive externalities on non-participants. Thus, a free-rider country, acting in a self-interest manner, can increase its own emissions and enjoy the benefits from the overall pollution reduction brought about by the coalition. In terms of stability, non-signatories' strong free-riding incentives imply a violation of the internal stability condition (3.50).

The second effect is related to the presence of trade and henceforth referred to as trade effect. In trade agreements, the coalition formation generates negative externalities on non-participants reducing their welfare (Yi, 1997). Thus, trade measures can be a key factor in increasing cooperation incentives. The intuition is that, if non-signatories expect trade barriers, they may have incentives to join IEAs. In this case, the external stability condition (3.51) is violated.

Even though in IEAs, members cannot exclude non-members from enjoying the benefits of a better global environment, in trade agreements, non-members can be excluded from enjoying the benefits of free trade. Free trade is not treated as a global public good, thus, when non-signatories express interest to cooperate (i.e. condition (3.51) is violated), they will not be admitted to the agreement if this action makes the existing members worse off. To that extent, we add one more condition, called admissibility condition. The admissibility condition takes the following form,

$$\mathcal{W}_s(s^*) > \mathcal{W}_s(s^* + 1), \tag{3.52}$$

and implies that even if external stability is violated and non-signatories wish to join, signatories oppose to the enlargement.

Considering the trade aspect in isolation, suppose that an agreement consisting of $n-1$ signatory countries is internally but not externally stable. That is, the last country has strong incentives to join the agreement as well. If the existing members admit the last country as a new member, they will gain tariff-free access to one new member country, but they grant the new member tariff-free access to $n-1$ countries. The new member must be better off, however, there is no guarantee that the existing members become better off (Yi, 1996). The admissibility condition is needed to ensure that existing members will admit a new member if they become better off by expanding the coalition.

In this context, stability is defined as follows:

Definition 11 *A GA of size s^* is stable if either,*

(i) $\mathcal{W}_s(s^*) \geq \mathcal{W}_{ns}(s^* - 1)$ *and* $\mathcal{W}_{ns}(s^*) \geq \mathcal{W}_s(s^* + 1)$ *or,*

(ii) $\mathcal{W}_s(s^*) \geq \mathcal{W}_{ns}(s^* - 1)$, $\mathcal{W}_{ns}(s^*) < \mathcal{W}_s(s^* + 1)$ *and* $\mathcal{W}_s(s^*) > \mathcal{W}_s(s^* + 1)$.

Furthermore, we want to point out that there is an important difference between environmental and trade agreements. In environmental agreements, countries can freely enjoy the benefits of a better environment free-riding on others efforts. Also, they can promote environmental quality by reducing their emissions without necessarily belonging to an agreement. However, in trade agreements, countries can benefit from free trade only if they are members of the agreement (i.e. benefits are excludable).

Solving analytically for the stability conditions under the two policies, i.e. tariffs and emission taxes, has proven impossible thus far.

3.6 Numerical analysis

In this section, we demonstrate a numerical analysis to study the model and to provide further intuitions. We are interested in examining the effect of trade on the stability and effectiveness of environmental agreements. Therefore, we focus on studying whether the formation of a GA can increase participation incentives, decrease global emissions and improve welfare relative to the corresponding outcomes of the basic model of the IEAs' literature.

3.6.1 The effect of trade on stability for $n = 10$

We use the following baseline parameter values: $n = 10$, $a = 1$, $b = 1$ and $\gamma = 0.045$.[8] Recall that parameter γ measures the impact of environmental damages

---

[8]Note that $\gamma = 0.045$ fails to satisfy the constraint (3.69) since $\gamma > 0.0433$. That is, in the IEA model the outcome is the global non-cooperation case, instead of the typical coalition of size 2.

to benefits due to emissions. In the IEA model, a stable agreement exists, though small, if that impact is low enough.[9] In the present analysis, we set a parameter value such that environmental damages have a more important effect on countries' welfare. Therefore, environmental pollution matters and countries apply strong enough environmental policies (emission taxes) to fight climate change. That is, we limit the range of the values of the parameter γ so that we generate interior equilibrium values for taxes for most coalition sizes. When the coalition size is expanded considerably, the equilibrium taxes of the non-signatories take a corner solution. In the present analysis, given that $n = 10$, $a = 1$ and $b = 1$, equilibrium taxes are mainly positive for $\gamma \geq 0.0415$. For that reason we choose $\gamma = 0.045$. Moreover, working with parameters that in the IEA model generates no stable coalition only strengthens our hypothesis that trade enhances cooperation.

Table 3.1 presents the production levels, net imports and consumption levels for each signatory and non-signatory country respectively.[10] The production and consumption levels for a signatory are given by the following equations, $e_s^p = e_s^d + (s-1)e_{ss}^X + (n-s)e_{sns}^X$ and $e_s^c = e_s^d + (s-1)e_{ss}^I + (n-s)e_{sns}^I$. On the other hand, the production and consumption levels for a non-signatory are given by the following equations, $e_{ns}^p = e_{ns}^d + (n-s-1)e_{nsns}^X + se_{nss}^X$ and $e_{ns}^c = e_{ns}^d + (n-s-1)e_{nsns}^I + se_{nss}^I$. The net imports for a signatory are equal to $(n-s)(e_{sns}^I - e_{sns}^X)$ while net imports for a non-signatory are equal to $s(e_{nss}^I - e_{nss}^X)$.[11]

[9] Parameter γ should satisfy the constraint (3.69).

[10] Values are rounded to four decimal places.

[11] Based on our notation, $e_{ss}^I = e_{ss}^X$ and $e_{nsns}^I = e_{nsns}^X$.

Table 3.1: Production, consumption and trade activities

	Signatory Country			Non-signatory Country		
s	Production	Net Imports	Consumption	Production	Net Imports	Consumption
1	–	–	–	0.5198	–	0.5198
2	1.3913	−0.7251	0.6661	0.4591	0.1813	0.6404
3	0.6643	−0.0213	0.6430	0.6246	0.0091	0.6337
4	0.5793	–	0.5793	0.6642	–	0.6642
5	0.5007	–	0.5007	0.7043	–	0.7043
6	0.4225	–	0.4225	0.7314	–	0.7314
7	0.3647	–	0.3647	0.6657	–	0.6657
8	0.3094	–	0.3094	0.5726	–	0.5726
9	0.2511	–	0.2511	0.5000	–	0.5000
10	0.1818	–	0.1818	–	–	–

The analysis shows that trade between signatories and non-signatories takes place only for small coalitions such that $s = 2$ and $s = 3$. In particular, signatories export while non-signatories import. For larger coalitions, net imports are zero (i.e. $e^I_{sns} = 0$ and $e^I_{nss} = 0$), however, trade still takes place but only among coalition members (i.e. $e^I_{ss} > 0$) and among non-members (i.e. $e^I_{nsns} > 0$).

Notice that, when members and non-members engage in trade activities, signatories report higher production and consumption levels. That is, the trade effect prevails. The inverse holds when they stop exchanging goods. In that case, signatories, controlled by the environmental policy, tend to gradually reduce their polluting activities.[12] However, non-signatories due to free-riding incentives increase theirs. For coalitions with size $s > 6$ we observe a decrease in non-signatories' production and consumption levels. The intuition is as follows: even though non-signatories increase the quantity they produce and consume domestically (i.e. e^d_{ns}) the total volume of trade among them gradually decreases since there are less countries outside the agreement.

Figure 3.1 displays the aggregate emissions, i.e. E (solid line). Additionally, we include in the graph the global emissions (dashed line) of the IEA model under the non-cooperative case i.e. $E^{nc} = 6.8966$.[13]

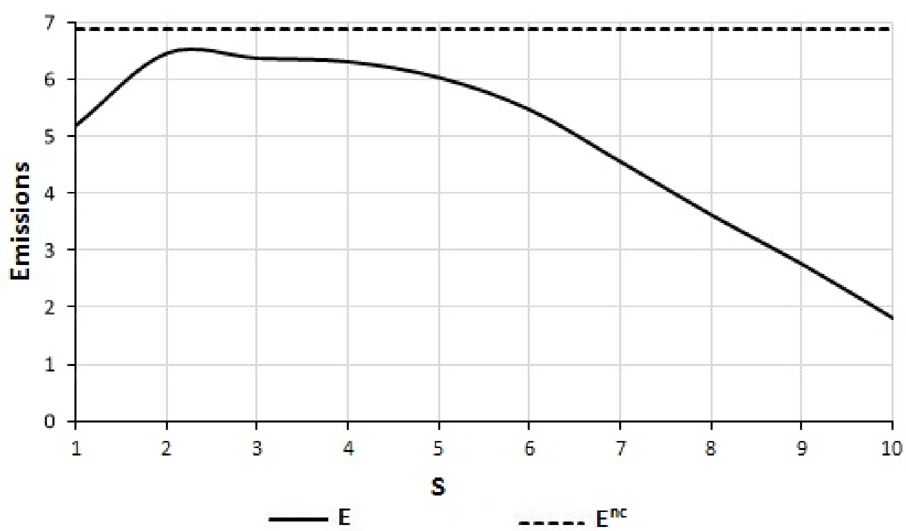


Figure 3.1: Aggregate emissions

[12]Signatories set a tax level that internalizes the full environmental cost of all coalition members. Thus, the larger is the coalition, the higher will be the emission tax.

[13]Global emissions in the non-cooperative case are calculated using equations (3.70).

Note that the dashed line does not denote the level of emissions per coalition, it denotes only the benchmark Nash equilibrium case. Results show that aggregate emissions decrease as we move to larger coalitions. Clearly, the environmental policy incites signatories to reduce significant their polluting activities. Moreover, as indicated in the graph, aggregate emissions in the GA model are lower than aggregate emissions in the IEA model. That is, the formation of a GA agreement can significantly improve on the basic model of the IEAs' literature in terms of environmental performance, especially when the coalition size increases.

The following Remark summarizes the aforementioned results.

Remark 12 *Regardless of stability, the interaction between trade and environment policies is essential to improve environmental protection.*

The following table, Table 3.2, presents the total welfare for each signatory country, i.e. $\mathcal{W}_s$, and non-signatory, i.e. $\mathcal{W}_{ns}$, country respectively, for any coalition size s. Additionally, we include in the table the global welfare defined by $\mathcal{W}_T = s\mathcal{W}_s + (n-s)\mathcal{W}_{ns}$.[14] When trade between signatories and non-signatories is present, the former are better off than the latter. Specifically, signatories receive higher welfare than non-signatories for coalitions with size $s = 2$ and $s = 3$. For those coalitions, the trade effect dominates the environmental effect. On the other hand, in the absence of trade, non-signatories become better off. Due to free-riding incentives, they report higher economic (polluting) activities than the coalition members

[14] Values are rounded to four decimal places.

receiving higher welfare. We want to point out that in trade agreements, the welfare of the non-signatories decreases in the size of the coalition (Yi, 1996). Moreover, notice that the expansion of the agreement improves the global welfare. That is, the global welfare is maximized under the grand coalition.[15]

Table 3.2: Welfare levels

s	W_s	W_{ns}	W_T
1	–	−0.2233	−2.2329
2	−0.4270	−0.5188	−5.0046
3	−0.4721	−0.4799	−4.7759
4	−0.4821	−0.4501	−4.6289
5	−0.4414	−0.3605	−4.0096
6	−0.3376	−0.2069	−2.8536
7	−0.1676	−0.0217	−1.2385
8	−0.0334	0.1138	−0.0395
9	0.0482	0.2036	0.6375
10	0.0909	–	0.9091

Remark 13 *Regardless of stability, the formation of GA improves welfare relative to the basic model of the IEAs' literature.*

Figure 3.2 depicts countries' welfare and is used to illustrate the effect of trade on stability. The welfare for the signatories, i.e. $\mathcal{W}_s(s)$, is depicted by the solid line and the welfare for the non-signatories, i.e. $\mathcal{W}_{ns}(s)$, is depicted by the dot-dashed line. Moreover, the welfare $\mathcal{W}_{ns}(s-1)$ is depicted by the dotted line and

[15]In the non-cooperative case of the IEA model, each country receives welfare $\mathcal{W}^{nc} = -0.6183$ (calculated using equation (3.71)). The global welfare is $\mathcal{W}_T^{nc} = -6.1831$.

represents the welfare for the non-signatories shifted by one. We use such line to represent graphically the internal stability condition as the vertical distance captures $\mathcal{W}_s(s) \lesseqgtr \mathcal{W}_{ns}(s-1)$.

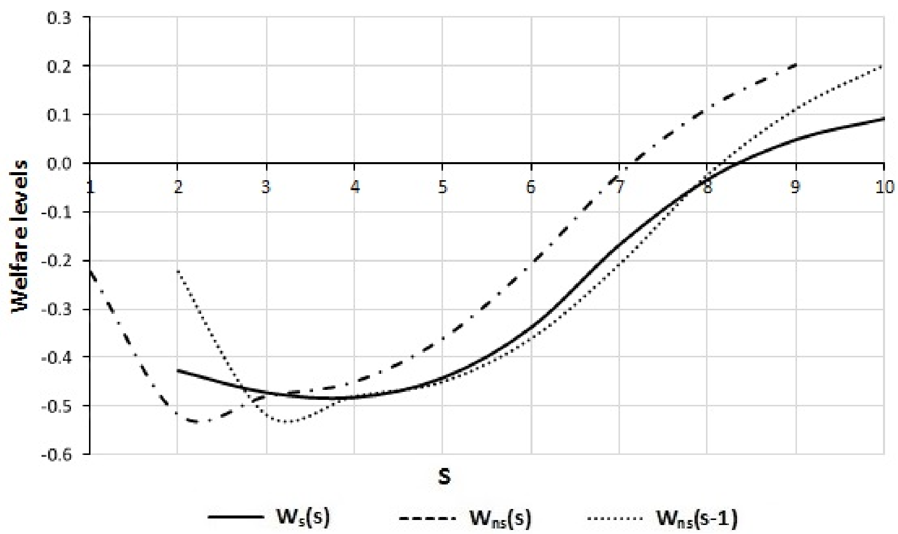


Figure 3.2: Welfare levels, $n = 10$

Signatories receive higher welfare than non-signatories when the trade effect prevails (the solid line is above the dot-dashed line), while the inverse holds when the environmental effect prevails (the dot-dashed line is above the solid line). The welfare level of non-members increases in the size of the coalition, that is, starting from the coalition of size two and gradually expanding the agreement makes non-members better off. The welfare level of members increases also in the size of the coalition but only for $s > 4$. As long as the trade affect prevails, global emissions are still high and as a consequence high environmental damages affect their welfare. For larger coalitions ($s > 4$) we observe a significant decrease in global emission level and thus the expansion of the agreement makes members better off.

As indicated in the graph, a stable agreement is achieved at $s^* = 7$. The agreement with seven countries is stable according to Definition: part (i), meaning that none of its participating countries has an incentive to withdraw (internal stability) and none of the non-participating countries has an incentive to further participate (external stability). To put it differently, at $s^* = 7$, the solid line in above the dotted line (i.e. $\mathcal{W}_s(7) > \mathcal{W}_{ns}(6)$). Therefore, internal stability is satisfied. Additionally, $s^* = 7$ is externally stable (i.e. $\mathcal{W}_{ns}(7) > \mathcal{W}_s(8)$) since the dotted line is above the solid line.

There exists also a small coalition that is also stable (Definition: part (i)) at $s = 3$ for the same reasoning as previously mentioned. However, it is important to note that the coalition $s = 3$ is not profitable since the welfare of the signatory countries is smaller than their welfare in the non-cooperative Nash equilibrium (Carraro and Siniscalco, 1993). Furthermore, note that the coalitions with sizes $s = 5$ and $s = 6$ fail to satisfy stability because they violate Definition: part (i) and (ii). Specifically, at those coalitions even though the internal stability is satisfied, the external stability is violated and also the admissibility condition is violated as well since $\mathcal{W}_s(5) < \mathcal{W}_s(6)$ and $\mathcal{W}_s(6) < \mathcal{W}_s(7)$ respectively. Hence, coalition members become better off by expanding the coalition.

The following Remark summarizes our findings.

Remark 14 *The size of a stable coalition increases when trade policies are included in the formation of an environmental agreement. Emissions are significantly lower and welfare higher at the stable coalition when compared to the corresponding outcomes of the IEA model.*

It is worth mentioning that by choosing a very low value for the parameter γ (for example $\gamma = 0.004$), such that emission taxes for non-signatories take a corner solution for any coalition size, we get similar results with Yi (1996) who examines the formation of customs unions.[16] The intuition of setting a very low value for the parameter γ is that environmental damages are not so severe for the countries and thus the environmental effect almost disappears from our model while the trade effect becomes very strong. In this case non-members have strong incentives to participate in a GA and benefit from free trade (i.e. external stability condition (3.51) is violated). However, admitting non-members to the agreement makes existing member worse off. The admissibility condition (3.52) holds for any coalition size greater than two. Thus, a stable coalition cannot have more than two members (Definition: part (*ii*)). Yi (1996) shows that the number of equilibrium customs unions in an unanimous regionalism game (a game where a customs union allows entry of a new member if and only if all existing members agree to admit the new member) is not greater than two.

[16] In his model the representative consumer in country i has a utility function similar to ours (assuming homogeneous goods). There are no environmental damages since he examines the formation of customs unions when countries trade with each other choosing their tariff levels.

3.6.2 The effect of trade on stability for larger n

In order to provide further implications of our study, we present two indicative examples with different parameter values and examine their effect on the derived results. In both cases, we use a graph, similar to the one presented in Figure 3.2, to illustrate how trade affects the stability of an agreement. We plot the welfare for the signatories $\mathcal{W}_s(s)$ (solid line), the welfare for the non-signatories $\mathcal{W}_{ns}(s)$ (dot-dashed line) and the welfare $\mathcal{W}_{ns}(s-1)$ (dotted line).

In the first numerical exercise, Example 1, we set $n = 15$, $a = 1$, $b = 1$ and $\gamma = 0.025$.[17] The welfare levels are presented in Figure 3.3. Trade between signatories and non-signatories takes place for coalitions with size $s = \{2, 3, 4\}$. In particular, the former export while the latter import. For those coalitions, members receive higher welfare than non-members (the solid line is above the dot-dashed line). For larger coalitions, such that $s > 4$, non-members become better off (the dot-dashed line is above the solid line). In this case, a stable agreement exists at $s^* = 11$ (Definition: part (i)). As indicated in the graph, at $s^* = 11$, the solid line in above the dotted line (i.e. $\mathcal{W}_s(11) > \mathcal{W}_{ns}(10)$). Thus, internal stability is satisfied. Moreover, $s^* = 11$ is externally stable (i.e. $\mathcal{W}_{ns}(7) > \mathcal{W}_s(8)$) since the dotted line is above the solid line. There exists also a small coalition (Definition: part (i)) that is internally and externally stable at $s = 5$, however, it is not profitable.

[17] Given that $n = 15$, $a = 1$ and $b = 1$, equilibrium taxes are mainly positive for $\gamma \geq 0.0216$.

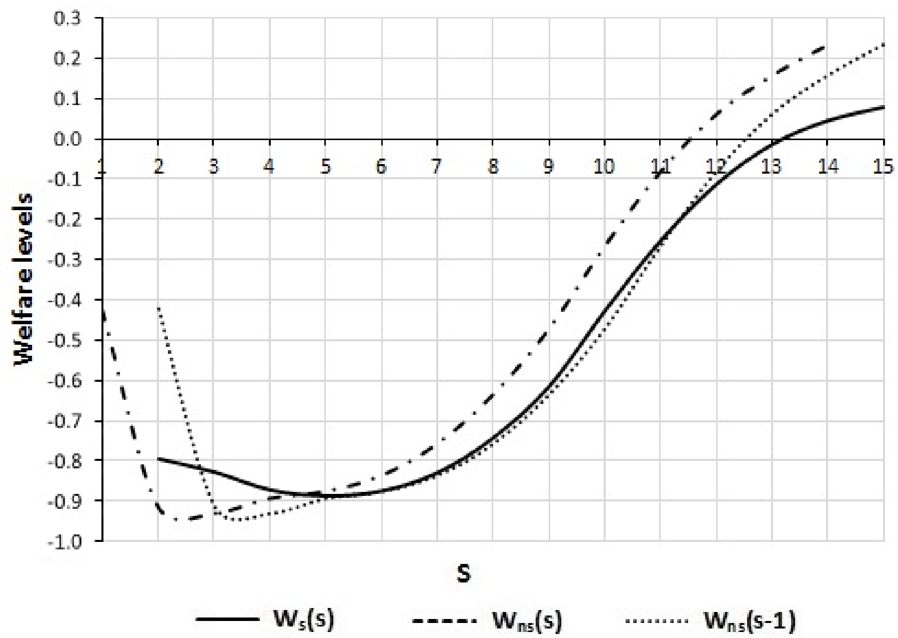


Figure 3.3: Welfare levels, $n = 15$

Example 2, presents the case where $n = 20$, $a = 1$, $b = 1$ and $\gamma = 0.015$.[18] Figure 3.4 illustrates the welfare levels. Signatories and non-signatories engage in trade activities for coalitions with size $s = \{2, 3, 4, 5\}$. In particular, the former export while the latter import. For those coalitions, members receive higher welfare than non-members. For larger coalitions such that $s > 5$, non-members become better off. In this case, a stable agreement is achieved at $s^* = 14$ (Definition: part (i)). There exists also a small coalition (Definition: part (i)) that is internally and externally stable at $s = 6$, however, it is not profitable.

[18] Given that $n = 20$, $a = 1$ and $b = 1$, equilibrium taxes are mainly positive for $\gamma \geq 0.013$.

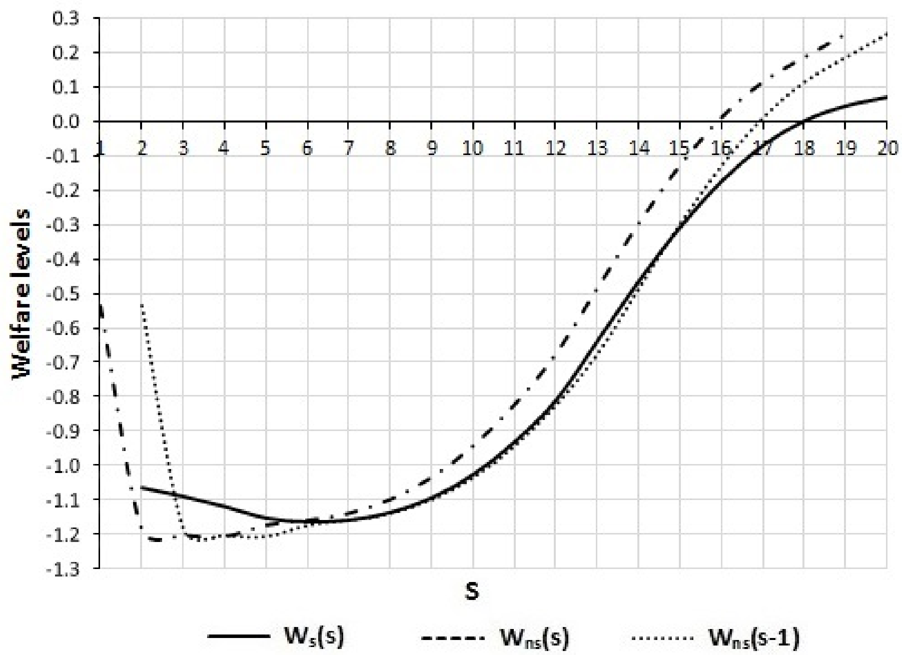


Figure 3.4: Welfare levels, $n = 20$

When we increase the number of countries (i.e. n), setting parameter values (i.e. γ) such that environmental pollution is important for them and so they take the necessary environmental policies to fight climate change, we find that the size of a stable agreement increases as well.[19] This result appears to contradict the main conclusion reached in the IEAs' literature stating that in a Cournot-IEA the maximum level of cooperation consists of two members, independently of the number of countries. However, the close to 70% participation seems to be quite robust against changes in n when other parameters are kept same.

[19] There are interior equilibrium values for taxes for most coalition sizes.

To summarize, results illustrate that trade along with environmental policies have an important impact on the stability and effectiveness of IEAs.

3.7 Conclusions

There has been considerable debate on the extent to which trade and environmental problems affect each other. Clearly, trade measures affect countries' production and consumption activities. Therefore, if these activities affect the environmental quality, trade will affect the environment. Similarly, environmental policies aiming to protect countries' environment influence the volume of trade.

The present paper examines the formation and stability of GAs. We extent the basic model of the IEAs' literature by letting homogeneous countries apply policy instruments such as emission taxes and tariffs in order to tackle the climate change problem and control trade. In this framework, countries are either members of a GA or outsiders. Each member (signatory country) has tariff-free access to other members' markets. On the other hand, each non-member (non-signatory country) pays tariffs on its imports to the other countries.

Results are optimistic when the IEA model is extended to incorporate trade. That is, the formation of an environmental agreement can be more successful when environmental policies are linked with trade policies. Countries have stronger incentives to cooperate and take the necessary measures to protect the environment. Thus, we can achieve larger stable agreements that reduce substantially aggregate emissions and improve welfare. Moreover, the analysis illustrates that the size of a stable agreement increases in the number of countries affected by the externalities.

The main limitation of our study is that the robustness of the derived results is not clear because analytical complexity requires resorting to numerical calculations.

To sum up, trade measures in IEAs can be an effective tool. Put in other words, the recommendation that stems out of this paper is that countries should not negotiate over environmental issues only. Rather, they should negotiate global agreements over at least two issues. The main concern is that trade policies applied in environmental agreements should be always compatible with the rules of the World Trade Organization (WTO) and its non-discrimination principle known as "most favoured nation treatment", which requires countries to grant equivalent treatment to the same products imported from any WTO member country (General Agreement on Tariffs and Trade (GATT)).

3.8 Appendices

3.8.1 Appendix A

Recall that, we define parameter γ as the ratio between environmental damages and benefits due to emissions. That is,

$$\gamma = \frac{c}{b}. \tag{3.53}$$

The reaction function for the equilibrium tariff τ_i (after imposing $\tau_i = \tau_j$ and $t_i = t_j$) is given by,

$$\tau_i = \frac{ab(\gamma n^2 + 3) + (n - 2 - \gamma n^2)t_i}{n + 7 + (n-1)n\gamma}. \tag{3.54}$$

The reaction function for the equilibrium tax t_i (after imposing $\tau_i = \tau_j$ and $t_i = t_j$) is given by,

$$t_i = \frac{ab(\gamma n^3 + (n-6)n + 4) - (n-1)(\gamma n^2 + 2n - 9)\tau_i}{(4n-5)n + 2 + \gamma n^3}. \tag{3.55}$$

3.8.2 Appendix B

The second order condition for the signatories' welfare maximization problem with respect to the tariff τ_s is satisfied,

$$\frac{\partial^2(sW_s)}{\partial \tau_s^2} = -(n-s)s\frac{n + 2 + (2s+3)s + (n-s)s^2\gamma}{(n+1)^2 b} < 0. \tag{3.56}$$

The second order condition for the signatories' welfare maximization problem with respect to the tax t_s is satisfied,

$$\frac{\partial^2(sW_s)}{\partial t_s^2} = -s^2\frac{2(n-s+1)n + (n^2\gamma - 1)s}{(n+1)^2 b} < 0. \tag{3.57}$$

The second order condition for the non-signatories' welfare maximization problem with respect to the tariff τ_{ns} is satisfied,

$$\frac{\partial^2 W_{ns}}{\partial \tau_{ns}^2} = -\frac{(n-2)n + 8s + 1 + (n-1)^2(n-s)\gamma}{(n+1)^2 b} < 0. \tag{3.58}$$

The second order condition for the non-signatories' welfare maximization problem with respect to the tax t_{ns} is satisfied,

$$\frac{\partial^2 W_s}{\partial t_{ns}^2} = -(n-s)\frac{n + (2n+1)s + (s+1)(n-s)n\gamma}{(n+1)^2 b} < 0. \tag{3.59}$$

3.8.3 Appendix C

The reaction function for the signatories' equilibrium tariff $\tau_s(t_s, \tau_{ns}, t_{ns})$ is given by,

$$\tau_s(t_s, \tau_{ns}, t_{ns}) = \frac{\begin{pmatrix} ab(1+(2+\gamma n^2)s) - ((n-s)ns\gamma - ((n-2(s+1))s-1))t_{ns} \\ -((n\gamma-2)s-1)st_s - (n-1)(n-s)s\gamma\tau_{ns} \end{pmatrix}}{\gamma(n-s)s^2 + (2s+3)s + n + 2}. \tag{3.60}$$

The reaction function for the signatories' equilibrium tax $t_s(\tau_s, \tau_{ns}, t_{ns})$ is given by,

$$t_s(\tau_s, \tau_{ns}, t_{ns}) = \frac{\begin{pmatrix} abn(n+1-(1+\gamma n^2)s) + (n-s)((\gamma n^2 s + n - s + (n-2s)n)t_{ns} \\ +((n-1)ns\gamma - 2(n+1-2s))\tau_{ns} - ((2-\gamma n)s+1)s\tau_s) \end{pmatrix}}{((2n+1)s - 2(n+1)n - n^2 s\gamma)s}. \tag{3.61}$$

The reaction function for the non-signatories' equilibrium tariff $\tau_{ns}(t_{ns}, \tau_s, t_s)$ is given by,

$$\tau_{ns}(t_{ns}, \tau_s, t_s) = \frac{\begin{pmatrix} ab((n-1)(\gamma n^2-1)+4s) - (n-s)((n-1)(1+\gamma n) - 4s)t_{ns} \\ -(3n+1-4s+(n-1)n\gamma)st_s - (n-1)(n-s)s\gamma\tau_s \end{pmatrix}}{(n-1)^2((n-s)\gamma+1) + 8s}. \tag{3.62}$$

The reaction function for the non-signatories' equilibrium tax $t_{ns}(\tau_{ns}, \tau_s, t_s)$ is given by,

$$t_{ns}(\tau_{ns}, \tau_s, t_s) = \frac{\left(\begin{array}{c} abn(s+1)(n(n-s)\gamma - 1) + ((n(n-2s)-s) - (n-s)(s+1)n\gamma)st_s \\ -(n-s)(n-1-4s+(n-1)(n-s)(s+1)\gamma)\tau_{ns} \\ -(s+1)(n-1-2s+(n-s)^2\gamma)s\tau_s \end{array}\right)}{(n-s)((n-s)(s+1)n\gamma + n + (2n+1)s)}. \tag{3.63}$$

3.8.4 Appendix D

The equilibrium levels of the tariffs, τ_s and τ_{ns}, for a signatory and non-signatory country respectively, are the following,

$$\tau_s = \frac{A_\tau}{D} \text{ and } \tau_{ns} = \frac{B_\tau}{D}, \tag{3.64}$$

where the expression A_τ is given by,

$A_\tau = ab((s+1)n^5 + (-s^2+2s+3)n^4 - (2s^2+s-3)n^3 - (s^3-4s^2+4s-1)n^2 + 2(-3s^2+5s+1)sn - 3(3s+1)s^2 + \gamma(sn^7 + (5s+1)n^6 + (-12s^2+3s+1)n^5 + (12s^3 - 11s^2+2s-1)n^4 + (8s^4+s^3-29s^2-4s-1)n^3 + (-5s^4+34s^3+53s^2+13s+1)sn^2 - (14s^3+22s^2+11s+1)s^2n + (3s^2+4s+1)s^3)),$

the expression B_τ is given by,

$B_\tau = ab(s(2n^4 + (s^2-s+6)n^3 + (-s^3+6s^2+5)n^2 + (-5s^3+6s^2+6s+3)n - (6s^2+5s+1)s) + \gamma(n^7 + 2(s^2+s+2)n^6 + (-7s^3-5s^2+s+5)n^5 + (4s^4+s^3-7s^2 - 4s+2)n^4 + (4s^4-2s^3-11s^2-18s-8)sn^3 + (-2s^5+13s^4+37s^3+37s^2+8s -$

$1)sn^2 - (5s^3 + 19s^2 + 20s + 4)s^3 n + (3s^2 + 4s + 1)s^4)),$

the expression D is given by,

$D = \gamma(n^7 + (3s^2 + 3s + 4)n^6 + (-5s^3 + 2s + 5)n^5 + (s^4 - 17s^3 - 25s^2 - 13s + 2)n^4 + (8s^4 + 18s^3 + 28s^2 + 11s - 8)sn^3 + (-8s^4 + 23s^3 + 17s^2 + 6s + 12)s^2n^2 + (2s^5 - 26s^4 - 14s^3 - 5s - 1)s^2 n + (6s^4 + 2s^3 - s^2 + 1)s^3) + n^6 + 2(s^2 + s + 2)n^5 - (4s^3 + s^2 + 2s - 5)n^4 + 2(s^4 - 2s^3 + 2s^2 - 4s + 1)n^3 + (5s^3 + 20s + 4)sn^2 + 2(-s^4 + 4s^3 + s + 2)sn - 3(2s^3 + s^2 + 1)s^2.$

The equilibrium levels of the taxes, t_s and t_{ns}, for a signatory and non-signatory country respectively, are the following,

$$t_s = \frac{A_t}{D} \text{ and } t_{ns} = \frac{B_t}{D}, \tag{3.65}$$

where the expression A_t is given by,

$A_t = -ab(n^5 + (s^2 + 1)n^4 + (-3s^2 + 5)sn^3 + (2s^4 - 2s^3 - 7s^2 - 1)n^2 + (2s^4 + 13s^3 + 14s^2 + s - 1)n + 2(-2s^3 + 4s + 1)s + \gamma(n^7 + (s^2 - 4s + 3)n^6 - 2(7s^3 + 5s^2 + 10s + 2)n^5 + (26s^4 + 27s^3 + 60s^2 + 27s - 4)n^4 - (18s^5 + 18s^4 + 97s^3 + 54s^2 - 17s - 3)n^3 + (4s^6 - 10s^5 + 64s^4 + 27s^3 - 35s^2 - 7s + 1)n^2 + (6s^5 - 30s^4 - 2s^3 + 27s^2 + 6s - 1)sn + 2(3s^3 - s^2 - 3s - 1)s^3)),$

the expression B_t is given by,

$B_t = -ab(n - s)(n^4 + (2s^2 + 3s + 4)n^3 + (-2s^3 + 3s^2 + 3s + 5)n^2 + (-4s^3 - 7s + 2)n - (2s^2 - 7s - 1)s - \gamma(n^5(3s + 1) + 2(3s^3 + 3s^2 + 5s + 2)n^4 - (12s^4 + 16s^3 + 27s^2 + 4s - 5)n^3 + (4s^5 - 3s^4 + 48s^3 + 31s^2 + 2)n^2 + (4s^4 - 20s^3 - 10s^2 - 3s - 1)sn - (s^2 - 4s - 1)s^2)).$

3.8.5 Appendix E

For comparison purposes, we present the solutions of the basic model where a Cournot-IEA consisting of two countries is the unique self-enforcing IEA. In particular, we lay out the solutions derived by Rubio and Casino (2001).

Total emissions are given by,

$$E^c = \frac{an}{1 + n\gamma + s(s-1)\gamma}. \tag{3.66}$$

Signatories receive welfare given by,

$$\mathcal{W}_s^c = \frac{1}{2}a^2 b\left(1 - \frac{n^2\gamma}{(1 + n\gamma + s(s-1)\gamma)^2}(s^2\gamma + 1)\right). \tag{3.67}$$

Non-signatories receive welfare given by,

$$\mathcal{W}_{ns}^c = \frac{1}{2}a^2 b\left(1 - \frac{n^2\gamma}{(1 + n\gamma + s(s-1)\gamma)^2}(\gamma + 1)\right). \tag{3.68}$$

A coalition consisting of two countries is the unique self-enforcing IEA if and only if parameter γ satisfies the following condition,

$$\gamma \leq \frac{1}{n - 4 + 2\sqrt{n^2 - 3n + 3}}. \tag{3.69}$$

In the non-cooperative case, the basic model gives the following solutions.

Total emissions are given by,

$$E^{nc} = \frac{an}{1 + n\gamma}. \tag{3.70}$$

Countries receive welfare given by,

$$\mathcal{W}^{nc} = \frac{a^2 b(1-(n-2)n\gamma)}{2(1+n\gamma)^2}. \quad (3.71)$$

3.8.6 Appendix F

The following list, presented in Table 3.3, includes the main variables used in the present study. Note that it is not exhaustive.

Table 3.3: List of selected notation

Notation	Explanation
e_i^d	Country i's quantity produced and consumed domestically.
e_s^d	Signatory's quantity produced and consumed domestically.
e_{ns}^d	Non-signatory's quantity produced and consumed domestically.
$e_{ij}^I(=e_{ji}^X)$	Quantity country i imports from country j.
$e_{ij}^X(=e_{ji}^I)$	Quantity country i exports to country j.
$e_{ss}^I(=e_{ss}^X)$	Signatory's imports from another signatory country.
$e_{sns}^I(=e_{nss}^X)$	Signatory's imports from a non-signatory country.
$e_{nss}^I(=e_{sns}^X)$	Non-signatory's imports from a signatory country.
$e_{nsns}^I(=e_{nsns}^X)$	Non-signatory's imports from another non-signatory country.
e_i^p	Country i's total production of the nonnumeraire good.
e_s^p	Signatory's total production of the nonnumeraire good.
e_{ns}^p	Non-signatory's total production of the nonnumeraire good.
e_i^c	Country i's total consumption of the nonnumeraire good.
e_s^c	Signatory's total consumption of the nonnumeraire good.
e_{ns}^c	Non-signatory's total consumption of the nonnumeraire good.

BIBLIOGRAPHY

1. ALMER, C. AND WINKLER, R. (2017). "Analysing the effectiveness of international environmental policies: The case of the Kyoto Protocol". *Journal of Environmental Economics and Management*, **82**, 125-151.

2. ASHEIM, G. B., FROYN, C. B., HOVI, J. AND MENZ, F. C. (2006). "Regional versus Global Cooperation for Climate Control". *Journal of Environmental Economics and Management*, **51(1)**, 93-109.

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

4. BARRETT, S. (1997). "The strategy of trade sanctions in international environmental agreements". *Resource and Energy Economics*, **19**, 345-361.

5. BARRETT, S. (1997). "Heterogeneous international environmental agreements". *Carraro, C. (Ed.), International Environmental Negotiations*, Edward Elgar, Cheltenham, UK, (Chapter 2).

6. BARRETT, S. (1999). "A Theory of Full International Cooperation". *Journal of Theoretical Politics*, **11(4)**, 519-541.

7. BARRETT, S. (2001). "International cooperation for sale". *European Economic Review*, **45(10)**, 1835-1850.

8. Barrett, S. (2005). "The Theory of International Environmental Agreements". *K. Mäler and J. Vincent, ed., Handbook of Environmental Economics*, **3**, 1458-1516.

9. Biancardi, M. (2010). "International Environmental Agreement: A Dynamic Model of Emissions Reduction". *Nonlinear Dynamics in Economics, Finance and Social Sciences*, Ed. by G. I. Bischi, C. Chiarella, and L. Gardini, Springer, 73-93.

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

11. Botteon, M. and Carraro, C. (1997). "Burden-sharing and Coalition Stability in Environmental Negotiations with Asymmetric Countries". *International Environmental Negotiations: Strategic Policy Issues*, Ed. by Carraro C, Edward Elgard, 26-55.

12. Botteon, M. and Carraro, C. (2001). "Environmental Coalitions with Heterogeneous Countries: Burden-sharing and Carbon Leakage". *Environmental Policy, International Agreements, and International Trade*, Ed. by A. Ulph, Oxford University Press, 26-55.

13. Calvo, E. and Rubio, S. (2013). "Dynamic Models of International Environmental Agreements: A Differential Game Approach". *International Review of Environmental and Resource Economics*, **6(4)**, 289-339.

14. CARRARO, C., EYCKMANS, E. AND FINUS, M. (2006). "Optimal transfers and participation decisions in international environmental agreements". *The Review of International Organizations,* **1**, 379-396.

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

16. CARRARO, C. AND SINISCALCO, D. (1998). "International environmental agreements: Incentives and political economy". *European Economic Review,* **42**, 561-572.

17. CHANDER, P. AND TULKENS, H. (1995). "A Core-theoretic Solution for the Design of Cooperative Agreements on Transfrontier Pollution". *International Tax and Public Finance*, **2(2)**, 279-293.

18. CHANDER, P. AND TULKENS, H. (1997). "The Core of an Economy with Multilateral Environmental Externalities". *International Journal of Game Theory*, **36**, 379-401.

19. 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.

20. CONCONI, P. AND PERRONI, C. (2002). "Issue linkage and issue tie-in in multilateral negotiations". *Journal of International Economics*, **57**, 423-447.

21. 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.

22. De Cara, S. and Rotillon, G. (2001). "Multi-Greenhouse Gas International Agreements". *Working paper, Agricultural Economics*, 2003-03.

23. de Zeeuw, A. (2008). "Dynamic effects on the stability of international environmental agreements". *Journal of Environmental Economics and Management,* **55**, 163-174.

24. de Zeeuw, A. (2015). "International environmental agreements". *Annual Review of Resource Economics*, **7**, 151-168.

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

26. Diamantoudi, E. and Sartzetakis, E. (2015). "International environmental agreements: coordinated action under foresight". *Economic Theory*, **59**, 527-546.

27. Diamantoudi, E. and Sartzetakis, E. (2017). "International Environmental Agreements - The Role of Foresight". *Environmental and Resource Economics*, in press, DOI: 10.1007/s10640-017-0148-1.

28. Eichner, T. and Pethig, R. (2013). "Self-enforcing environmental agreements and international trade". *Journal of Public Economics*, **102**, 37–50.

29. Eichner, T. and Pethig, R. (2014). "Is trade liberalization conducive to the formation of climate coalitions?". *International Tax and Public Finance*, **22**, 932–955.

30. Eichner, T. and Pethig, R. (2015). "Self-enforcing international environmental agreements and trade: taxes versus caps". *Oxford Economic Papers*, **67(4)**, 897–917.

31. Eyckmans, E. and Finus, M. (2004). "An Almost Ideal Sharing Scheme for Coalition Games with Externalities". *Working paper*, No. 14/2014.

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

33. Finus, M. and McGinty, M. (2017). "The Anti-Paradox of Cooperation: Diversity Pays!". *Working paper*, No. 40/15.

34. Fernandez-Amador, O., F.-Francois, J. and Tomberger, P. (2016). "Carbon dioxide emissions and international trade at the turn of the millennium". *Ecological Economics*, **125**, 14-26.

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

36. Gallo, C., Faccilongo, N. and La Sala, P. (2018). "Clustering analysis of environmental emissions: A study on Kyoto Protocol's impact on member countries". *Journal of Cleaner Production*, **172**, 3685-3703.

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

38. IPCC (2104). "Climate Change 2014: Synthesis Report". Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

39. Kesternich, M. (2016). "Minimum participation rules in international environmental agreements: empirical evidence from a survey among delegates in international climate negotiations". *Applied Economics*, **48(12)**, 1047-1065.

40. Kuriyama, A. and Abe, N. (2018). "Ex-post assessment of the Kyoto Protocol – quantification of CO_2 mitigation impact in both Annex B and non-Annex B countries". *Applied Energy*, **220**, 286-295.

41. Lessmann, K., Marschinski, R. and Edenhofer, O. (2009). "The effects of tariffs on coalition formation in a dynamic global warming game". *Economic Modelling*, **26**, 641-649.

42. Lise, W. and Tol, R. (2004). "Attainability of International Environmental Agreements as a Social Situation". *International Environmental Agreements*, **4(3)**, 253-277.

43. McGinty, M. (2007). "International environmental agreements among asymmetric nations". *Oxford Economic Papers*, **59**, 45-62

44. Nordhaus, W. (2015). "Climate Clubs: Overcoming Free-riding in International Climate Policy". *American Economic Review*, **105(4)**, 1339-1370.

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

46. Oxfam International (2015). "Game-changers in the Paris climate deal: What is needed to ensure a new agreement helps those on the front lines of climate change". *Series: Oxfam Media Briefings.*

47. Oxfam International (2015). "Extreme Carbon Inequality". *Series: Oxfam Media Briefings.*

48. Petrakis, E. and Xepapadeas, A. (1996). "Environmental consciousness and moral hazard in international agreements to protect the environment". *Journal of Public Economics*, **18**, 51-68.

49. Pavlona, Y. and de Zeeuw, A. (2013). "Asymmetries in International Environmental Agreements". *Environment and Development Economics*, **18**, 51-68.

50. Rubio, J. S. and Casino, B. (2001). "International Cooperation in Pollution Control". *Working paper*, Serie AD 2001-21.

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

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

53. Weikard, H. P. (2009). "Cartel Stability under Optimal Sharing Rule". *Manchester School*, **77**, 575-593.

54. Yi, S.-S. (1996). "Endogenous formation of customs unions under imperfect competition: open regionalism is good". *Journal of International Economics*, **41**, 153-177.

55. Yi, S.-S. (1997). "Stable Coalition Structures with Externalities". *Games and Economic Behavior*, **20**, 201-237.