Stability of international climate treaties

The importance of heterogeneity

Runa Haave Andersson

Thesis for the degree
Master of Economic Theory and Econometrics

Department of Economics

UNIVERSITY OF OSLO

January 10, 2013

STABILITY OF INTERNATIONAL CLIMATE TREATIES

THE IMPORTANCE OF HETEROGENEITY

2013

Stability of international climate treaties – the importance of heterogeneity

Runa Haave Andersson

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

Preface

I want to thank my main supervisor, Karine Nyborg, for all her time spent discussing every part of this thesis with me and for thoroughly reading all previous versions of it. Her contributions have been invaluable.

I also want to thank Statistics Norway for letting me use an office and all their resources. In particular I want to thank my supervisor Bjart Holtsmark for introducing me to this topic and for his excellent guidance throughout this process.

Finally, I want to thank Dritan Osmani for quickly providing me with the necessary links to the model description for the FUND model.

Table of contents

List of tables

List of figures

1 Introduction and summary

Climate change presents serious global risks that require global collective action in response; see for example IPCC (2007) and the Stern Review (2006). The challenge is to make global cooperation work. Climate change is therefore a central topic in international politics today, and has been for the last two decades. Nevertheless, little is happening on a global scale to reduce emissions and stabilise the concentration of greenhouse gases in the atmosphere.

From a theoretical point of view this is not surprising. Existing literature on international climate treaties indicates that international cooperation will not succeed in solving the problem of climate change. A main reason for this is that it is much more profitable for each individual country to stay out of an agreement and let the other countries do all the work. Everybody wants to be a free-rider.

This thesis, however, presents a somewhat optimistic view. I will use a linear benefit/quadratic cost model to analyse possible coalitions between 16 independent countries. My approach is based on an analysis by Bjart Holtsmark (2013), where he uses a set of spreadsheets containing a calculation procedure of his own design to look at the stability properties of all possible coalitions between 16 parties (to be specified below). Holtsmark (2013, p. 3) bases his calculations "on non-cooperative game theory and the concept of stable coalitions as originally conceived by d'Aspremont et al. (1983)". He examines possible coalitions between 16 world regions, where the parameter values were estimated by Osmani and Tol (2009).

As Holtsmark (2013) mentions, single countries represented by their government are the relevant parties in international negotiations. A model of international climate treaties should therefore have individual countries, not groups of countries, as units. When there are 16 potential parties to a treaty, there are 65 519 possible coalitions with at least two members. With a larger number of parties, the number of possible coalitions is also much larger. Ideally the model should include all countries that are affected by emissions of climate gases, i.e. all the countries in the world. The task of examining all possible coalitions with at least two members for every country in the world is, however, too massive for this thesis.

I have applied Holtsmark's spreadsheets as basis for the results in this thesis. I had to choose 16 countries to include in the model and then estimate the necessary parameters based on the estimations of Osmani and Tol. These parameter values allow me to get numerical results on the possible coalitions from Holtsmark's spreadsheets.

The countries that have the highest emissions are also the countries that can contribute most to total reduction in emissions on a global scale. I have therefore in this thesis decided to include the European Union in addition to the 15 countries with the highest emissions of CO_2.[1] Those countries are: China, the United States of America, the Russian Federation, India, Japan, Canada, the Republic of Korea, the Islamic Republic of Iran, Mexico, South Africa, Australia, Saudi Arabia, Brazil, Ukraine and Indonesia.

The European Union is a union of several individual countries. However, since the European Union can make decisions that are binding for its member states (Barrett, 2003) and the union negotiate internationally as a single unit, it is natural to include EU as a single party in the model.

In chapter 2 I will briefly present the parts of environmental economic theory that are of relevance to understanding the model I will use in this thesis. The model itself will be presented in chapter 3, where I will solve it first for two identical countries, then for 16 identical countries. I finally introduce the model with 16 heterogeneous countries in section 3.3.

In chapter 4 I go through the methodological and theoretical foundation for my own parameter calculations. These are, as previously noted, based on parameters presented in the paper "Toward Farsightedly Stable International Environmental Agreements" by Osmani and Tol (2009). The choice of the FUND model and the modification of these parameters are inspired by Holtsmark (2013). At the end of the chapter I present my final parameter calculations.

In chapter 5 I present the numerical results from using my parameters in Holtsmark's spreadsheets (2013) and discuss some key coalitions. The spreadsheets are based on the model formulation presented in section 3.3 and calculates the payoffs and abatement level for each country in each of the 65 520 possible coalitions (including the outcome where no countries cooperate). The spreadsheets also calculate the global payoff and the global abatement for each coalition and rate the coalitions according to global payoff and global

---

[1] In 2005 (the World Bank, 2005a)

abatement – relative to global payoff and global abatement in the two extremes where no one cooperates and where everyone cooperates.

This is a brief summary of my findings from chapter 5:

The spreadsheets calculate which coalitions are internally stable and which coalitions are potentially internally stable. As expected, there are relatively few internally stable coalitions. Some internally stable coalitions do, however, have a large number of members. These large coalitions unfortunately do not perform well in terms of relative global abatement (or in terms of relative global payoff).

The majority of the 65 520 possible outcomes are potentially internally stable when side payments are introduced. While not even the best performing internally stable coalitions have high relative global abatement, a large number of the potentially internally stable do. By allowing for side payments in this thesis, we massively increase the gains of potential cooperation. The main conclusion of this thesis is that when we have a model with heterogeneous countries and allow for side payments we can get quite close to the social optimum (as defined later in the thesis).

While a coalition also needs to be externally stable in order to be self-enforcing (according to the definition of self-enforceability introduced later in this thesis) I do not calculate external stability. This is mainly because calculating this is outside the scope of this thesis. I could also argue that internal stability is the most interesting concept; while external stability ensures that no outsider wants to be part of the treaty, internal stability ensures that no member wants to leave the treaty. Forcing a sovereign state to stay on as a member of a treaty is not possible according to international law. However, forming an exclusive treaty with restrictive membership is possible.

Chapter 6 is the conclusion. In addition to briefly commenting on the main results from chapter 4 and 5, I also point out some interesting questions I have not had time to answer in this thesis.

2 Some basic theory

2.1 Pollution

All economic activity takes place within the natural environment, i.e. the earth and its atmosphere. Resources extracted from the environment are used as inputs in production or directly in consumption (Perman et al., 2011, ch. 2). Both consumption and production give rise to waste products, or residuals, that are released back to the environment. The emission of residuals that are perceived to be harmful is called pollution (Perman et al., 2011, ch. 2).

According to *the materials balance principle* matter can neither be created nor destroyed (Perman et al., 2011, ch. 2). Matter extracted from the environment is simply transformed to make it more useful to humans, and the mass not contained in these products is discharged as waste and returned to the environment. Residuals are thus the unintended by-products of the transformation process, and pollutants are residuals that harm the environment (Perman et al., 2011, ch. 2).

The mixing of the pollutant refers to how it spreads out after it is emitted (Perman et al., 2011, ch. 5). A pollutant is uniformly mixing if it spreads out equally over the relevant space; all that matters is how much is emitted in total, not the source or the location of the emissions. One example is the emission of most green-house gases (Perman et al., 2011, ch. 5).

The harmful effect of pollution is often referred to as the *damage* from pollution. The damage depends either on the stock or the flow of the pollutant (or a mix of the two). The flow of a pollutant is the rate at which it is discharged into the environment and the stock is the accumulated concentration rate of the pollutant in the environment (Perman et al., 2011, ch. 5). Damage that arises only from the flow of a pollutant will instantly drop to zero when the emission drops to zero. Damage that depends on the stock of pollution is affected by emissions only insofar as they add to the existing stock (Perman et al., 2011, ch. 5). Green-house gases are typically stock pollutants.

Pollution is per definition harmful and this, in isolation, indicates that pollution should be avoided. Assuming that rational producers have initially chosen production techniques that maximize their profits and that pollution is an unintended by-product of the production process, the producers either have to reduce production or invest in cleansing systems and/or less polluting production technologies in order to reduce emissions (Perman et al., 2011, ch. 5). Reducing emissions leads to less output and/or increased costs; this is the cost of abatement. Thus, there is a trade-off between the damage from pollution and the cost of reducing emissions (Perman et al., 2011, ch. 5).

2.2 Market failure

An allocation of resources is inefficient if it is possible to increase the welfare of one or more individuals without reducing the welfare of anyone else. Such inefficiencies often stem from some kind of market failure; when one or more of the conditions for a well-functioning market are not fulfilled.

Externalities are examples of such a market failure. "By definition, an externality arises where the action of one party affects the set of outcomes attainable by another, and where this effect is *not* taken into account by the party undertaking the action" (Barrett, 2003, p. 75). An externality can originate in both production and consumption, and can have either positive or negative effects (Perman et al., 2011, ch. 4). The emission of green-house gases is a harmful externality. An externality is internalized if the generator(s) is forced to account for the unintended consequences; *every* activity has to have a price (Sandler, 1997).

The Coase Theorem states that if property rights are well defined private bargaining between rational, self-interested individuals leads to efficient outcomes even in the presence of externalities (Stiglitz, 2006). There are, however, several examples of situations where property rights are difficult to define.

The atmosphere is a global common property; no single state or individual can make binding decisions regarding the use of the atmosphere (Barrett, 2003, ch. 5). This implies that the Coase Theorem does not provide us with a solution to the problem of green-house gas emissions.

We could consider the atmosphere as a *public good*. The existence of public goods is another reason for market failure. Two properties characterise pure public goods: non-rivalry and non-excludability (Perman et al., 2011, ch. 4). Once a public good is made available, no one (payers and non-payers alike) can be prevented from consuming the good (non-excludability). One individual's consumption of the public good does not, in any way, hinder the consumption by others of that same unit of the good (non-rivalry) (Sandler, 1997). These properties undermine the private incentive to provide and maintain the goods, and they are thus not readily provided by the market. By letting others provide the good (and pay the cost), it is possible to *free-ride* on the efforts of others (Sandler, 1997). When one can enjoy the good just as much anyway, why bother paying for it? If everyone reasons this way, the good may not be provided. It may be necessary for some sort of government to intervene in order to supply the goods, but in different ways and to a varying degree, depending on the specific characteristics of the good itself. As explained by Sandler (1997, p. 43): "The class of public goods is large and can vary in terms of the degree of non-rivalry and the extent of non-excludability."

Abatement of green-house gas emissions can also be seen as a global public good; every unit of abatement is to the benefit of everyone, not only those who pay for the abatement.

There can be no externalities and no public goods if a market is to be defined as well-functioning. In addition there must be perfect competition (no single actor can influence the prices) and perfect information (Stiglitz, 2006). Real markets are most often not well-functioning, and efficient outcomes are often not realized (Perman et al., 2011, ch. 4).

2.3 International law

Why is there a basic difference in the difficulties of solving environmental problems at national and international levels?

By different means, democratic or otherwise, a state can make general rules, or laws, to solve market failure at the national level. The government has the power to assign and enforce property rights, redistribute income, regulate externalities and provide (or maintain) public goods (Sandler, 1997). Conflicts can be brought before a national court, where the decisions of the court are legally binding and can be enforced, seeing as all citizens are bound by the laws of their nation (Barrett, 2003).

Market failures at the international level, such as transborder externalities (externalities that arise across borders, and that affect several states), can only be corrected by voluntary cooperation between the affected parties (this and the following paragraphs are based on discussions in Barrett (2003)). International law must be recognised by all states the laws are meant to apply to, and a dispute can only be brought before the International Court of Justice (the judicial organ of the UN) if all parties agree to it. Even if the parties do agree to go to court, no one has the authority to enforce the court's decision.

Custom dictates that states are sovereign equals; they are entitled to administer their own affairs without foreign interference and have equal right to act internationally. States accept some limitations on their sovereignty, knowing that their actions may inadvertently infringe upon the rights of other states. Perhaps more importantly, the actions of other states may conflict with their own right not to be interfered with and they accept constraints themselves because they want other states to accept the same limitations.

Customary law is a tradition-bound institution; it is not created, but evolves naturally. It's a pattern of behaviour that arises from repeated interaction between states and leads to expectations about how states should act internationally). Customary law is enforced by informal mechanisms, there are no formal mechanisms to force compliance and punish misbehaviour. Any state that deviates from existing custom must be prepared to defend themselves to other states that see this behaviour as being unlawful. A deviation might also encourage deviation in general, by eroding international cooperation. Unless they have good reasons not to, most states simply obey custom and behave as other states expect them to, even if it in isolated cases might not be in their own best interest.

Treaties are used to facilitate transnational collective action (actions that require cooperation between at least two states to further the interests of them all) (Sandler, 1997), and are agreements between states specifying how they should behave. They are created through international deliberation and negotiations, and specify a set of rights and obligations (Barrett, 2003). Each state comes to the negotiating table intent on getting the best possible deal for themselves, and while customary law dictates that all members should adhere to the treaty, it also dictates that each state is free to choose whether or not to sign the agreement (Barrett, 2003).

2.4 Game theory

This thesis applies game theory; the concept is briefly introduced in this section (for a more thorough description of the subject, see Gibbons (1992) or Watson (2002)). Game theory can be used to gain some understanding about the underlying motivations of the participants in international negotiations. States guard their autonomy, and international cooperation is loose as a result of this. They can therefore be compared to "game players that must choose their strategies based on their beliefs about the likely choices of others" (Sandler, 1997, p. 51).

A player acts *unilaterally* if he is acting on his own behalf, without consulting anyone else. He acts in his own *self-interest* if he cares only about his own payoff, without regard for the payoff other players get (Barrett, 2003). When several actions give the same payoff, a player is *indifferent* between them. He will not prefer one above the other.

According to Watson (2002, p. 5): “[g]ames are formal descriptions of strategic settings”. *Formal* implies a logically consistent and mathematically precise structure. When the actions of one player are interdependent with the actions of other players and all players recognise this interdependence, the players behave *strategically* (Sandler, 1997). One player will try to anticipate what the other does, and act upon that. The other player anticipates that the first player will anticipate and acts according to this. Of course, this has already been anticipated by the first player, something the other also anticipates... And so on. A *strategy* is a set of strategic responses to every possible action of others. A player’s *best response* is the strategy that maximizes his payoff, given his belief about the action of the other players (Watson, 2002). If one specific strategy always yields a higher payoff than the others, no matter what the other players do, that strategy is a *dominant strategy* (Watson, 2002).

When a game has been played, an outcome is realized. Given that all players have played from their best responses, an outcome is strategically stable if no player will unilaterally want to deviate from their chosen strategy. Such a strategically stable outcome is called a *Nash equilibrium* (Watson, 2002). A game can have no, one or several Nash equilibrium outcomes. A Nash equilibrium is, however, not necessarily a *Pareto efficient* outcome (Watson, 2002). For an outcome to be Pareto efficient, all possible *Pareto improvements* must have been made. A Pareto improvement is possible when one player can be made better off without making any other player worse off. In some cases the players would all be better off if they all acted differently. If a Pareto efficient outcome is not a Nash equilibrium, it will not be realized as long as the players act unilaterally. That is because at least one of the players will do better by changing their position (given that the other players do not change theirs). To realize the Pareto efficient outcome the players then have to be able to coordinate upon and commit to a unified strategy.

There are, according to Sandler (1997), seventy-eight distinct 2 x 2 (a game with 2 players who have 2 choices each) game structures. The structure does not depend on specific payoffs, but on the ordering of the payoffs.[2] The most famous game is the *Prisoners' Dilemma*. The story goes like this (Watson, 2002):[3]

Two criminals are taken into questioning by the police, suspected of committing a crime together. They are questioned separately. The police make them both an offer. If only one of them confesses, he will get a prison sentence of 1 year while the other gets a sentence of 2.5 years. If both confess, they both get a sentence of 2 years. If none of them confesses, they will be charged with a minor offence and sentenced to 1.5 years each. Their choices are illustrated in table (2.1). Each cell shows the sentence for each criminal whether they confess or keep silent, given the choice of the other criminal. The first number in each cell is criminal 1's sentence and the second is criminal 2's sentence.

		Criminal 2			
		Confess		*Keep silent*	
Criminal 1	*Confess*	**2**	**2**	1	2.5
	Keep silent	2.5	1	**1.5**	**1.5**

Table 2.1 – The Prisoners' Dilemma

The first number in each cell is criminal 1's sentence and the second is criminal 2's sentence.

Based on theory presented in Watson (2002)

To confess while the other keeps silent is the individually best outcome for each criminal. The next best outcome is when both keep silent. The third best outcome is when both confess, and the individually worst outcome for each criminal is to keep silent while the other confesses. The collectively best outcome, where their total prison sentence is minimized (at 3 years), is when both keep silent. This is a Pareto efficient outcome; it will not be possible to make one criminal better off without making the other worse off. It is not, however, a Nash equilibrium. Given that the other keeps silent, both criminals will want to deviate and confess, hoping to achieve the individually best outcome. There is *one* Nash equilibrium in this game and that is the outcome where both confess; the collectively worst outcome with a total prison sentence of 4 years. Given that the other criminal confesses, they can do no better than to confess

---

[2] See Sandler (1997, p. 28) for a more thorough discussion.

[3] The general story is found in Watson on p. 31-*32*, though I use different numbers.

themselves. To confess is a dominant strategy; no matter what the other criminal does, the criminals can do no better than to confess. The Nash equilibrium is the *only* outcome in this game that is not Pareto efficient.

This is the characteristic structure of a Prisoners' Dilemma; a game with a Pareto inefficient Nash equilibrium where one action is a dominant strategy for both players, but where both players would be better off if they were able to coordinate and commit to the other strategy. The Prisoners' Dilemma is routinely used to describe environmental problems because so many environmental problems seem to be examples of this game structure (Perman et al., 2011, ch. 9).

2.5 International climate treaties

Solving a transboundary environmental problem requires international cooperation (Hoel, 2005). By designing a suitable agreement it should be possible to make every country better off than they would be if they acted unilaterally (Hoel, 2005). When states unilaterally make their optimal choice in the trade-off between environmental damage and the cost of reducing the damage, they will only consider their own costs and damages, taking the actions of other states as given. Simple actions may sometimes change incentive structures, so that self-interested actors, who would not otherwise do so, behave in a way that promotes the common good (Sandler, 1997, p. 25). According to Barrett (2003, p. 355) "[t]he principal task of a treaty is to restructure incentives." Unfortunately, nations are, according to Sandler (1997, p. 51), "calculating entities that serve their own interests." Reaching an agreement is difficult, since the states are different and will be affected differently by an agreement (Hoel, 2005).

Free-riding is a hindrance to the success of international climate treaties (IEAs). Many transboundary environmental problems are uniformly mixing, so that it does not matter where the problem originates. Climate change is a problem of this sort. It does not matter where the emission of climate gases takes place. Only the total stock of pollutants matter. Similarly, it does not matter where the abatement takes place. If a state manages to stay outside an agreement, it can enjoy the benefits from the reduced emissions of the cooperating parties, without having to bear any of the costs of reducing emissions themselves (Hoel, 2005). When everyone wants to be a free-rider, no one is willing to committ to a treaty. This can be a huge problem even if all states are better off with the agreement than without (Hoel, 2005).

To make self-interested states agree on a treaty and comply to its commitments it has to be profitable for them to do so (Finus, 2004). It must also be *self-enforcing*. That is, it must change the incentives of the parties in such a way that they do not want to unilaterally deviate from the agreement. To be self-enforcing an agreement must, according to Barrett (2003, p. xiv), be individually rational, collectively rational and fair. The threats and promises stated in the treaty have to be *credible*, i.e. all parties must believe that threats and promises are feasible and that they will be carried out if necessary (Barrett, 2003).[4] It is *individually rational* if the parties to the treaty have incentives to comply with and not withdraw from the agreement, given the choices of all other states. Furthermore, non-parties should have incentives neither to join the agreement nor change their present behaviour, given the choices of other states (Barrett, 2003). If it is not possible to gain collectively by changing the treaty, the agreement is *collectively rational* (Barrett, 2003). Scott Barrett (2003, p. xiv) claims that an agreement also needs to be *fair* in order to gain and keep the consent of all relevant parties.

The efficiency of an agreement can be measured against estimated abatement levels in the absence of a treaty (Finus, 2004), what is often called the *business as usual* (BaU) level of emissions. Large coalitions may not necessarily be more efficient than small coalitions, since for small coalitions it may be easier to agree on ambitious abatement levels and enforce the treaty (Finus, 2004). If the abatement target (the level of abatement they want to reach) specified in the IEA equals the BaU level, participation and compliance will not be a problem (Finus, 2004). It is not difficult for a state to agree to do something it would have done anyway. An ambitious treaty is more likely to specify targets that forces states to abate more than they would have done unilaterally, making it more likely that they will not sign the treaty at all. When designing a treaty one is often faced with a trade-off between the depth (how ambitious the treaty is) and the breadth (how many participants there are) of cooperation (Barrett, 2003, ch. 14).

Finus (2004) makes a distinction between two types of IEA-models: Membership models and compliance models. Within the membership models he further distinguishes between three

---

[4] See Watson (2002) for a more precise definition of credibility

sub-groups; models using cooperative game theory, models using non-cooperative game theory and 'new coalition theory'.[5]

The following discussion of the different IEA-models is, unless otherwise stated, based on Finus (2004).

Membership models are focused on coalition formation and the stability of membership. In this framework the central question is whether or not an agent is a party to an agreement. In the simplest form this is simply a decision of whether to abate or pollute. Free-riding in these models is not to be part of an agreement, i.e. to pollute. It is assumed that when an agent becomes a party to a treaty he will comply with its directives.

States exhibit cooperative behaviour when they cooperate and make agreements on how they should act. Their behaviour is non-cooperative when they make decisions unilaterally; maximizing their own payoffs conditional on an expectation of how the others will act (Perman et al., 2011, ch. 9). According to Perman et al. (2011, p. 285) "[a] widely accepted tenet of non-cooperative game theory is that dominant strategies are selected where they exist."

The earliest studies on membership models were rooted in cooperative game theory. The focus in these models is on the stability of the *grand coalition* and the socially optimal level of abatement. This can be interpreted as the situation where all potential parties form a coalition and jointly maximize their aggregate payoffs. The agents then have to decide whether or not to be party to this specific treaty. The gain from cooperation is the aggregate payoff from cooperation that exceeds the payoff from not having the coalition. The coalition is *stable* if each member's share of the gains from cooperation is high enough, so that they will not profit from not being members of the coalition. Side payments (see definition below) may be used to further induce participation. These models focus solely on first-best outcomes (the socially optimal outcome), and are not able to explain the sub-optimal treaties we see in real life.

Other studies on membership models also make use of non-cooperative game theory. The payoff structure in these models is static, meaning that there is only one time period (in which decisions are made simultaneously and the final outcome is determined). An individual payoff

---

[5] These will be further explained below.

is assigned to every agent for every possible coalition. Each coalition is assumed to maximize aggregate payoff conditional on actual number of members. Non-members maximize their own individual payoffs. Agents "cooperate with their coalition but behave non-cooperatively against outsiders" (Finus, 2004, p. 97). As a result, equilibrium is often a situation with lower than socially optimal abatement and inefficient coalition structures. The main question is: Which coalition structure can be sustained as an equilibrium? The concepts internal and external stability are central to answering this question.

A coalition is internally stable if no member is better off by no longer being a member and externally stable if no non-member is better off by becoming a member. Deviation, which in this framework means to be a non-member, is tempting to individual agents if it is profitable. When a member leaves the coalition, the remaining members re-optimize their aggregated payoff, conditional on the remaining number of members. The 'punishment' for leaving is then that the remaining coalition members increase their emissions somewhat. Leaving a coalition is profitable if the private gain from leaving exceeds the private damage from the increased emissions resulting from the coalition increasing their emissions. When a new member joins a coalition the reverse happens: The coalition members re-optimize their aggregated payoff (and decrease their emissions). Joining a coalition is profitable if the resulting decrease in damages is larger than the cost of abating.

When there are many potential parties and only the total abatement matters, the contribution of one state matters little in the aggregate. When many states are involved and the environmental damages are high compared to the benefits of emissions, the gains of cooperation are high (Finus, 2004). But this is also when the incentive to free-ride is large (Barrett, 2003, ch. 13). Thus; cooperation is harder to sustain the greater the need for cooperation is (Barrett, 2003, ch. 7). The main result is that for homogeneous agents, only small coalitions are stable, and for heterogeneous agents (without side payments) the coalitions accomplish even less in terms of emission reduction (Finus, 2004).

Pavlova and de Zeeuw (2012) look at coalitions when there are two types of countries; one type with higher benefits of emissions and lower damages than the other. Their conclusion is that while asymmetries among states allow for much larger coalitions in terms of members; total emission reductions are actually lower in these large coalitions than in small coalitions among only the countries with low benefits of emissions and high damages.

A coalition is potentially internally stable if the gains of cooperation are large enough to, with the use of side payments, compensate the members so that no member becomes better off by no longer being a member (Holtsmark, 2013). Side payment can be seen as incentive payment to make an agent agree to doing something it would not otherwise do, they can be helpful in sustaining cooperation and in supporting a fair outcome (Barrett, 2003, ch. 13). The payments can be made either in cash or in kind [6] (Finus, 2004) and the promise of side payments has to be credible. Side payments are only helpful when the agents are heterogeneous (not equal).[7] When there are strong asymmetries (differences) among the players, side payments are able to sustain a vastly superior outcome compared to the situation without side payments (Barrett, 2003, ch. 13). On the downside, while side payments do make recipients more inclined to become members, they make the donors 'worse off' by lowering their payoff, making them less likely to want to stay members (Barrett, 2003, ch. 13).

The models using non-cooperative game theory are able to look at several possible coalitions, not only the socially optimal one. Still, they are limited to looking at only one coalition at a time, and are not able to evaluate a situation where there exist multiple coalitions at once. This subject is investigated in the sub-group of membership models Finus names 'new coalition theory' (see Finus, 2004, p. 120).

Compliance models are the second type of IEA models (the following synopsis is based entirely on Finus (2004)). These models take membership as given. Free-riding in this context is to be a member of a treaty but not comply with the agreed abatement. Compliance "is only interesting in a non-cooperative game theoretical setting where there is a time lag between violation and discovery through other participants" (Finus, 2004, p. 125). The outcome is stable if incentives to free-ride are deterred by threats to punish deviations. Sanctions often have a negative effect also on those carrying out the punishment, and the coordination of sanctions between the remaining members may be costly and time-consuming, making it hard to see the threat of punishment as credible. For more information on compliance models, see Finus (2004, p. 125).

---

[6] In kind payments take the form of a gift of commodities rather than cash (Handmark, Inc., 2011)

[7] See Barrett (2003, ch. 13.2) for an explanation of why it does not make sense to look at side payments for homogeneous (identical) countries.

3 The model

In this chapter I present the model I will use in the rest of this thesis, using my own calculations. It is a model of international negotiation on the reduction of greenhouse gases. The motivation for looking at this model is found in Holtsmark (2013) and has been discussed in the introduction to the thesis (chapter 1). Scott Barrett (2003, ch. 7 & 13) describes a similar negotiation model.[8]

Emissions of CO_2 and other climate gases harm the atmosphere. The damages from pollution depend on the stock of pollutants. Emissions only matter in so far as they add to the existing stock. The emission of greenhouse gases is an example of uniformly mixing pollution. The location of emissions does not matter; an increase in the stock of pollutants only depends on total emissions.

In my model there is no distinction between the accumulation of and the emission of pollution. That is because the model is static; all decisions are made within one time period. The model does not consider changes in the stock of pollution over time.

The model I will use belongs to the group of models Finus (2004) calls membership models. The payoff function is the linear benefit/quadratic cost function found in Barrett (2006, p. 22):

$$\pi_i = b_i \sum_{j=1}^{N} q_j - \frac{c_i}{2} q_i^2 \qquad i, j = 1, .., N \qquad (3.1)$$

There are N countries, π_i is country i's payoff and q_i is country i's abatement. $b_i \sum_{j=1}^{N} q_j$ is country i's benefit from abatement and $\frac{c_i}{2} q_i^2$ is country i's cost of abatement. It is clear from this equation that while a country's benefit depends on the sum of all abatement (global abatement level), its cost depends only on the abatement in country i.

The benefit of abatement can be interpreted as damages avoided when emissions are reduced. I assume, for simplicity, that the damages avoided are (and can be precisely) measured in a monetary unit. If the benefit parameter, b_i, is positive, emissions are harmful for country i.

The cost of abatement is simply the cost of reducing emissions (see discussion in chapter 2.1), measured in a monetary unit. I assume that the cost parameter, c_i, is strictly positive. In other words, I assume that abatement has a positive cost for country i.

---

[8] Barrett uses a linear payoff function in this book.

Supplying abatement is privately profitable for country *i* as long as the total cost of abatement for the country is lower than the country's total benefit from that abatement. The abatement of country *i* is collectively profitable as long as the total cost of abatement for country *i* is lower than the global benefit (the sum of all individual countries' benefits) of that abatement. Assuming that the global benefit of abatement is larger than the individual benefit for country *i*, unilaterally determined abatement will fall short of the abatement level that is best for the global society.

I assume that each country behaving unilaterally maximize its individual payoff, taking the other countries' abatement as given. The privately optimal level of abatement (q_i^*) for country *i* is found by equating *i*'s marginal benefit and marginal cost. This is what is often referred to as the business as usual (BaU) level of abatement.

$$\max_{q_i} \pi_i \quad \Longrightarrow \quad b_i = c_i q_i \quad \Longrightarrow$$

$$q_i^* = \frac{b_i}{c_i} \tag{3.2}$$

b_i is *i*'s marginal benefit and $c_i q_i$ is *i*'s marginal cost. The marginal benefit can be interpreted as the increase in benefits when abatement is increased by one unit. The marginal cost is the increase in costs as abatement increases by one unit.

If the marginal cost for country *i* is lower than the marginal benefit for country *i*, the increase in benefits more than makes up for the increase in costs for country *i* following an increase in abatement. Country *i*'s payoff increases with abatement until the privately optimal level of abatement is reached. Beyond this point, *i*'s marginal benefit is smaller than *i*'s marginal cost, and a further increase in abatement reduces *i*'s payoff.

The payoff function for country *i* can be rewritten as:

$$\pi_i = b_i \sum_{j=1}^{N-1} q_{j \neq i} + \left(b_i q_i - \frac{c_i}{2} q_i^2\right) \qquad i = 1, \dots, N \tag{3.3}$$

The first term in equation (3.3) is strictly increasing in $q_{j \neq i}$, while the term in the brackets is decreasing in q_i for $q_i > \frac{b_i}{c_i}$ given that b_i and c_i are positive. This means that the payoff function is increasing with respect to the abatement of others, but is decreasing in own abatement above and beyond the privately optimal abatement level. Each country will prefer a situation where they themselves abate at the privately optimal level, while the others abate more.

We can make sure that the privately optimal abatement level is the only maximum of π_i with respect to q_i by looking at the second order condition (Sydsæter, 2000):

$$\pi_i''(q_i) = -c_i < 0 \quad \text{for all possible } q_i \text{ since } \; c_i > 0$$

We then know that equation (3.2) is the maximum of equation (3.1), which is strictly concave with respect to q_i. The payoff function is increasing in q_i for $q_i < q_i^*$ and decreasing in q_i for $q_i > q_i^*$, confirming the statement above (Sydsæter, 2000).[9]

A country's benefit depends on the abatement level in the other countries. Assuming that global welfare is given by the sum of the individual countries' payoff, the socially optimal level of abatement (q_i^{**}) is found by maximizing the global payoff (the sum of the all individual countries' payoff):

$$\max_{q_i} \sum_{i=1}^{N} \pi_i = \sum_{i=1}^{N} b_i \sum_{j=1}^{N} q_j - \sum_{i=1}^{N} \frac{c_i}{2} q_i^2 \implies \tag{3.4}$$

$$q_i^{**} = \frac{\sum_{i=1}^{N} b_i}{c_i} \tag{3.5}$$

Since the sum of strictly concave functions is a strictly concave function (Sydsæter, 2000), it follows that (3.5) is the maximum. It also follows that global payoff increases in q_i for $q_i < q_i^{**}$ and decreases in q_i for $q_i > q_i^{**}$.

To arrive at the socially optimal solution, there must exist some mechanism by which each country is forced to take the *global* benefit of one unit of abatement into account. It is obvious that as long as the sum of all individual benefits $(\sum_{i=1}^{N} b_i)$ is positive, the socially optimal level of abatement is larger than the privately optimal level of abatement.[10] As long as countries act unilaterally, emission reductions will be below the social optimum.

---

[9] $\pi_i'(q_i) \geq 0$ for $q_i \leq q_i^*$ and $\pi_i'(q_i) \leq 0$ for $q_i \geq q_i^*$ where q_i^* is a global maximum point (Sydsæter, 2000)

[10] $q_i^{**} = \frac{\sum_{i=1}^{N} b_i}{c_i} > q_i^* = \frac{b_i}{c_i} \qquad if \sum_{i=1}^{N} b_i \; and \; c_i > 0$

3.1 Two identical countries

In this section I will present the model assuming there are only two, identical countries. Abatement in country 1 is q_1, abatement in country 2 is q_2, making global abatement $q_1 + q_2$. Assume also that q_i is an integer from 0 to 3. The payoff for each country has the linear benefit/quadratic cost structure described above. Since the countries are identical they have the same costs and benefits. The value of the parameters b and c does not matter for this analysis, so I will follow Holtsmark (2013) and assume that $b = c = 1$. The payoffs for the two countries are then:

$$\pi_i = (q_1 + q_2) - \frac{1}{2} q_i^2 \qquad i = 1,2$$

The different payoffs corresponding to different abatement levels are presented in table 1. Each cell shows the payoff of each country for given values of q_1 and q_2. The first number in each cell is country 1's payoff and the second is country 2's payoff.

		Country 2							
		$q_2 = 0$		$q_2 = 1$		$q_2 = 2$		$q_2 = 3$	
Country 1	$q_1 = 0$	0	0	1	0,5	2	0	3	-1,5
	$q_1 = 1$	0,5	1	**1,5**	**1,5**	2,5	1	3,5	-0,5
	$q_1 = 2$	0	2	1	2,5	**2**	**2**	3	0,5
	$q_1 = 3$	-1,5	3	-0,5	3,5	0,5	3	1,5	1,5

Table 3.1 – An abatement game between two countries

Source: Holtsmark (2013)

The BaU level of abatement (see eq. (3.2)) is $q_i^* = 1$. This is the abatement each would choose if they act unilaterally. The socially optimal level of abatement is $q_i^{**} = 2$ (see eq. (3.5). This is the abatement they would choose if they cooperate – where the global payoff is maximized.

The choice, for each country, between 1 and 2 units of abatement is a classic Prisoners' Dilemma.[11] Abatement below the BaU level will never be chosen, since each country's payoff is increasing with respect to its own abatement below this point, making it profitable to increase abatement. Total payoff is increasing with country *i*'s abatement until the socially optimal level of abatement is reached, then decreases with further abatement in country *i*. Abatement above the socially optimal level will not be chosen if the countries are able to cooperate.

Both countries could do no better than to stick with the BaU level of abatement, regardless of what the other country does, hence the BaU abatement level is a Nash equilibrium. It is, however, not Pareto efficient. It is possible to make both countries better off if both choose the socially optimal level of abatement. Unfortunately, the BaU level of abatement is their dominant strategy. It is clear from table 3.1 that no matter what the other country does, both countries can do no better than to choose the BaU level of abatement. If they make an agreement to choose the socially optimal abatement level, both countries would want to deviate and choose the BaU level of abatement instead. To reach the socially optimal outcome, the countries must be able to negotiate and commit to an agreement. Since there is no mechanism to enforce international cooperation, such an agreement has to be self-enforcing. It has to change incentives so as to make it desirable not to deviate from the social optimum. An agreement between several countries is described in the next section.

3.2 16 individual countries

Now, assume that we have N = 16 identical countries with the payoff structure defined in section 3.1, where $b = c = 1$:

$$\pi_i = \sum_{j=1}^{16} q_j - \frac{1}{2} q_i^2 \qquad i = 1, \dots, 16 \tag{3.6}$$

From equation (3.2) I find that the BaU level of abatement is $q_i^* = 1$, giving each country a payoff of 15.5.[12] This is the level of abatement each country will choose if they make decisions unilaterally. Global payoff is then 248. The socially optimal level of abatement is

---

[11] The inner part of table (3.1) is in fact the mirror image of table (2.1) from chapter 2.

[12] Assuming that all 16 countries abate at the BaU level and inserting this into equation (3.5).

found by using equation (3.5); $q_i^{**} = \frac{Nb_i}{c_i} = 16$. Assuming that all countries abate at the socially optimal level, each country then receives a payoff of 128, making the global payoff 2048. This is clearly a much better solution for everyone. Unfortunately, the payoff from being a free rider (enjoying high abatement from others but abating at the privately optimal level (BaU) themselves) is 240.5 given that the other 15 countries still abate at the socially optimal level. This is clearly a much better payoff than what they get from abating 16 units like the rest.

We know from equation (3.3) and the subsequent discussion that q_i^* = 1 is country i's dominant strategy, because this is the abatement that maximizes country i's payoff function, irrespective of other countries' choices. In other words, no matter what the other countries do, it is best for country i to stick to the BaU level of abatement. Since this is true for all countries, the only Nash equilibrium is the outcome where every country abates only 1 unit each.

To arrive at a collectively better result, the countries can make an agreement. Since the incentive to free ride is present, the agreement has to be self-enforcing. The social optimum, where all 16 countries cooperate and maximize their joint payoff is, as we already know from the discussion above, not a stable solution. The incentive to free ride is too large.

Keeping in mind that global payoff increases with country i's abatement as long as $q_i < q_i^{**}$, it would be interesting to see if some sort of agreement could lead to a stable outcome that is better than the BaU solution. Assuming that countries behave cooperatively towards other members if they are part of an agreement and non-cooperatively towards those who are not makes this a non-cooperative membership model (Finus, 2004). Coalition members maximize their joint payoff, taking the behaviour of non-members as given. Non-members make unilateral decisions, maximizing their own payoff.

This can be seen as a three stage game (Barrett, 2003, ch. 7), where the countries at stage 1 decide whether or not to be part of an agreement. k is then the number of members in an agreement.[13] In stage 2 the coalition members chose their abatement level, taking the number of members and non-members as given, while in stage 3 the non-members unilaterally choose their own abatement level, also taking the number of members and non-members as given.

---

[13] k is an integer from 1 to 16.

Since the payoff structure is common knowledge, all countries know the abatement levels and payoffs that follow from every possible coalition size. It therefore makes sense to begin by looking at stage 3 first, then at stage 2 and finally at stage 1, taking the resulting abatement levels as given.

We already know that the BaU abatement level will be chosen by countries acting unilaterally. It does not matter how large the coalition is; although the payoff for non-members is increasing in *k* it is always best to abate only 1 unit. Thus, the non-members will choose $q_i^* = 1$ no matter what.

In stage 2 the coalition members decide which abatement level will be chosen for every coalition size. They maximize their joint payoff, taking the emission of non-members as given. The optimal abatement for each coalition member is $q_i^k = k$.[14]

Table 3.2 depicts the payoff for each member ($\pi^m(k)$) and non-member ($\pi^n(k)$) of every possible coalition size (*k*): [15]

k	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
$\pi^m(k)$	-	16	17,5	20	23,5	28	33,5	40	47,5	56	65,5	76	87,5	100	113,5	128
$\pi^n(k)$	15,5	17,5	21,5	27,5	35,5	45,5	57,5	71,5	87,5	105,5	125,5	147,5	171,5	197,5	225,5	-
Global payoff	248	277	332	410	508	623	752	892	1040	1193	1348	1502	1652	1795	1928	2048
Global abatement	16	18	22	28	36	46	58	72	88	106	126	148	172	198	226	256

Table 3.2 – The payoff for members (*m*) and non-members (*n*) and global payoff and abatement when there is a coalition with *k* members

See footnote 14 and 15 for calculation of $\pi^m(k)$ and $\pi^n(k)$. Global payoff is the sum of the payoff for members and non-members. Global abatement is the sum of abatement for members and non-members.

It is easy to see that non-members have a greater payoff than coalition members, and that the payoffs increase with *k*. Total abatement also increases with *k*. This model says nothing about which countries cooperate and which do not. The exact identity of members and non-members is irrelevant to the total payoff and total abatement when the countries are identical.

---

[14] $q_i^k = \frac{\sum_{i=1}^k b_i}{c_i} = \frac{kb_i}{c_i} = k \qquad k \in (1, 16)$

[15] We have that $\pi^m(k) = \sum_{m=1}^k q_m^k + \sum_{n=(k+1)}^{16} q_n^* - \frac{1}{2}(q_m^k)^2 = \frac{1}{2}k^2 - k + 16$

and that $\pi^n(k) = \sum_{m=1}^k q_m^k + \sum_{n=(k+1)}^{16} q_n^* - \frac{1}{2}(q_n^*)^2 = k^2 - k + \frac{31}{2}$

The identities are not, however, irrelevant to the individual countries themselves, since the countries that get to free ride have a much higher individual payoff than those who cooperate.

All that remains is to find out how many countries will choose to become coalition members. The agreement must be self-enforcing in the sense that no members should be better off by not being members. That means that there must be internal stability. Non-members must be better off by staying out of the agreement, i.e. there must be external stability. In terms of the payoff functions, this means that:

$$\pi^m(k) - \pi^n(k-1) \geq 0 \qquad (Internal\ stability)^{16} \qquad (3.7)$$

$$\pi^n(k) - \pi^m(k+1) \geq 0 \qquad (External\ stability)^{17} \qquad (3.8)$$

Figure 3.1 depicts the left-hand sides of the two stability equations ((3.7) and (3.8)). The equations are solved using the payoff calculations from table 3.2 above.

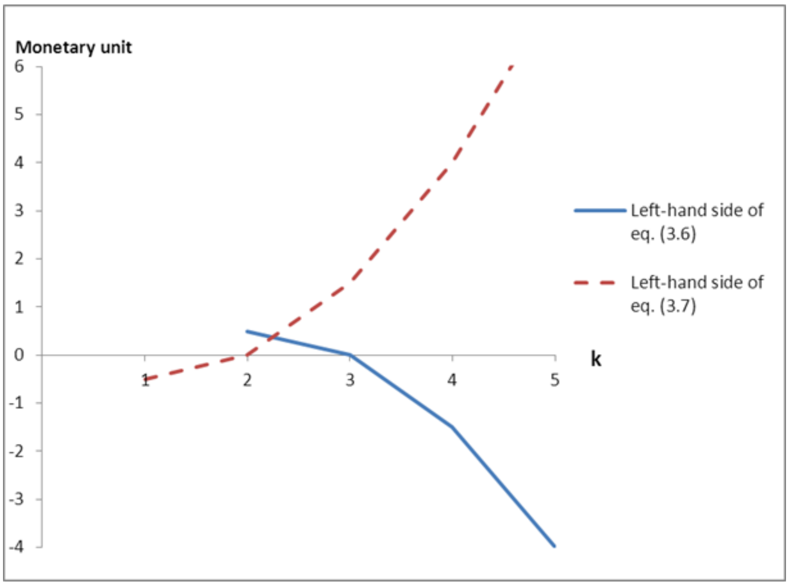

Figure 3.1 – The stability of a coalition with *k* members
A coalition is both internally and externally stable when both curves are ≥ 0.
Source: Payoff calculations from table 3.2

The internal stability condition (eq. (3.7)) holds for $k = 2$ and $k = 3$ and does not hold for k > 3. (Eq. (3.7) is not relevant for $k < 2$). This implies that the solutions $k = 2$ and $k = 3$ are the only internally stable solutions, see dashed line in figure 3.1.

---

[16] $\pi^m(k) - \pi^n(k-1) = \frac{1}{2}k^2 - k + 16 - (k-1)^2 + (k-1) - \frac{31}{2}$

[17] $\pi^n(k) - \pi^m(k+1) = k^2 - k + \frac{31}{2} - \frac{1}{2}(k+1)^2 + (k+1) - 16$

The external stability condition (eq. (3.8)) is binding for $k = 2$ and holds for $k > 2$, but does not hold for $k = 1$. Note that when $k = 1$ there is no cooperation and the countries behave unilaterally, abating at the BaU level. We therefore see that the BaU solution is not externally stable. However; all other solutions – i.e. where $k \geq 2$ – are externally stable.

A coalition is both internally and externally stable only when $k = 2$ or $k = 3$.

Although there are 16 countries and everyone has to cooperate to achieve the socially optimal solution, the largest coalition that is both internally and externally stable has only three members. Total payoff when $k = 3$ is 332. Although this is higher than the total payoff in BaU of 248, it is no way near as large as the socially optimal payoff of 2048. In terms of achieving higher abatement, more cooperation is better. Total abatement is only 16 units when $k = 1$. When $k = 3$ total abatement is slightly higher; 22 units. This is far below the socially optimal level of total abatement; 256 units. This rather depressing result, that we only get small coalitions with relatively low abatement, is a well-known result in theory of international climate treaties (see section 2.5). Still, total abatement *is* higher with cooperation than without.

3.3 Heterogeneous countries

In this section I will briefly present the model when we have 16 *heterogeneous countries*, expanding on the discussion in Holtsmark (2013).[18] That the countries are heterogeneous means that they are not identical, but have their own country specific costs and benefits. That means they have individual values of the parameters *b* and *c*. This is, in a sense, a more realistic assumption, since states really are different.

When states are different and *there are no side payments*, it is no longer obvious that a coalition will maximize its joint payoff. While the sum of payoffs is indeed maximized, some states may have much lower individual gains from cooperation than others.

Many different abatement rules could be negotiated, making some countries better off and some countries worse off. One alternative rule is that the members of a coalition agree to abate the same percentage of their existing emissions. One of the main problems of

---

[18] Here, Holtsmark looks at 16 heterogeneous world regions, but points out that single countries are in fact the relevant units.

international climate negotiations is getting all states to agree on a specific plan of action (Barrett, 2003).

If we allow for side payments, it is more realistic that the coalition will choose to maximize its joint payoff. The total gain of cooperation is then as large as it can be, and can be divided among the coalition members in any way they please.

In the following I will assume that the members of a coalition do maximize their joint payoff, both with and without side payments. I will first consider the situation without side payments.

Country *i*'s payoff function is identical to equation (3.1), with $N = 16$:

$$\pi_i = b_i \sum_{j=1}^{16} q_j - \frac{c_i}{2} q_i^2 \qquad i, j = 1, .., 16 \tag{3.9}$$

The BaU level of abatement (see eq. (3.2)) for country i is $q_i^* = \frac{b_i}{c_i}$. This depends only on the costs and benefits of country i. We know from the discussion at the beginning of this chapter that this is country *i*'s best response (the abatement level that maximizes *i*'s individual payoff) no matter what the other countries do.

The socially optimal level of abatement (see eq. (3.5)) for country i is $q_i^{**} = \frac{\sum_{i=1}^{16} b_i}{c_i}$. This depends on the costs of country *i* and all countries' benefit from abatement. Because of the concavity of the global payoff function (eq. (3.4)) with respect to q_i global payoff decreases with further abatement from country *i*, no matter what the other countries do.

Like the case with identical countries, the BaU level of abatement is the only Nash equilibrium when the countries make decisions unilaterally. It is therefore interesting to see if some sort of cooperation (and an outcome which is better for all countries than the BaU solution) is self-enforcing (see section 2.5). As with the identical countries, such an agreement must be both internally and externally stable.

When the countries have different payoff functions, the identity of the members and non-members of a coalition is relevant for aggregate payoffs. There are 65 519 different hypothetical coalitions with at least two members when we look at coalitions between 16 heterogeneous countries.[19]

Let K denote any possible coalition and M be the set of countries that are members of coalition K. The members of coalition K maximize their joint payoff with respect to q_i. q_i^K is then the abatement level for member i of coalition K:

$$\max_{q_i} \sum_{i \in M} \pi_i = (\sum_{i \in M} b_i)(\sum_{j \in M} q_j) - \sum_{i \in M} \frac{c_i}{2} q_i^2 \implies$$

$$q_i^K = \frac{\sum_{i \in M} b_i}{c_i} \qquad (3.10)$$

D is the set of non-members of coalition K. Non-members abate at the BaU level (q_i^* from eq. (3.2)).

$\Pi_i(K)$ is the payoff for country i provided that there exists a coalition K. Country i is a member of coalition K if $i \in M$ but is not a member of coalition K if $i \in D$:

$$\Pi_i(K) = b_i \left(\sum\nolimits_{j \in M} (q_j^K) + \sum\nolimits_{j \in D} (q_j^*) \right) - \frac{1}{2} \left(q_i(K) \right)^2$$

$$\text{where } q_i(K) = \begin{cases} \frac{\sum_{i \in M} b_i}{c_i} = q_i^k & if \quad i \in M \\ \frac{b_i}{c_i} = q_i^* & if \quad i \in D \end{cases}$$

Coalition K is internally and externally stable if these conditions hold:

$$\Pi_i(K) - \Pi_i(K \backslash i) \geq 0 \quad for\ i \in M \qquad (Internal\ stability) \qquad (3.11)$$

$$\Pi_i(K) - \Pi_i(K \cup i) \geq 0 \quad for\ i \in D \qquad (External\ stability) \qquad (3.12)$$

$\Pi_i(K \backslash i)$ is the payoff for member i if it withdraws from coalition K and $\Pi_i(K \cup i)$ is the payoff of non-member i if it joins coalition K.

---

[19] With n countries there are 2^n possible subsets of countries, including the empty set. By subtracting the empty set (no countries in the coalition) and the n single-country coalitions we get 2^n - (n + 1) coalitions with at least two members. When we have 16 countries we get: $2^{16} - (16 + 1) = 65\ 519$

The coalition is internally stable if country i's payoff (when $i \in M$) from being a member of the coalition is greater than (or equal to) the payoff country i gets from being a non-member of coalition $K \setminus i$; a coalition consisting of all the members of coalition *K* except for country i itself. This must hold for all members of coalition *K*.

Coalition *K* is externally stable if country i (when $i \in D$) gets a higher payoff from being a non-member of coalition *K* than from being a member of coalition $K \cup i$; a coalition with all the original members of coalition *K* plus country *i* itself. This must hold for all non-members of coalition *K*.

It is also interesting to investigate whether some sort of side payment can be made to induce countries to become members of a coalition they would not otherwise be members of. A coalition K is potentially internally stable if it can be made internally stable when side payments are allowed.

$$\sum_{i \in M} (\Pi_i(K) - \Pi_i(K \setminus i)) \geq 0 \quad i \in M \qquad (Potential\ internal\ stability) \qquad (3.13)$$

This condition states that the sum of all members' net gains from membership must be positive. Thus, coalition *K* is potentially internally stable even if country *i*'s internal stability condition does not hold, as long as at least one other member of the coalition gains enough from being part of the coalition to make up for *i*'s loss. The countries that gain can make side payments to the countries that lose to entice them to be members of the coalition.

The payoffs, abatement level and stability of the 65 519 possible coalitions depend on the parameter values of all the 16 countries. The parameter values will be discussed in the next chapter. The results of the model will then be presented and discussed in chapter 5.

The results are obtained by using the spreadsheets developed by Holtsmark (2013). The spreadsheets incorporate the equations from the model for heterogeneous countries specified above, calculating country i's payoff and abatement level, total payoff and abatement and the individual countries' gains from participation in each possible coalition (see the left hand side of equation (3.11)). This allows us to analyse the internal stability and potential internal stability of each coalition.

Both Holtsmark and I have checked a number of coalitions with manual calculation to confirm the accuracy of the spreadsheets.

4 The parameters

To be able to analyse the 65 519 coalitions and check whether they are internally, externally and potentially internally stable I need to know the numerical values of b_i and c_i for all countries i. The values should reflect, as precisely as possible, the real costs and benefits for each country. I need to find a simple, reasonable and consistent approximation of the parameters.

Anything from cleaner air to the preservation of an endangered species or limiting the global rise in temperature to avoid damaging changes in agriculture and a rising sea level can be benefits from reduced green-house gas emissions. For some countries there may even be positive consequences of global warming, such as better conditions for agriculture due to increased temperatures at higher altitudes. This would translate into a negative benefit parameter. My interpretation of the benefit parameter (b_i) is that it is country i's marginal willingness to pay for one unit of reduction in emissions, i.e. one unit of abatement.

The cost of abatement depends on what kind of technology a country uses, how energy efficient they are, how essential the polluting technology is to the country's economy, etc. I assume that the countries abate where abatement is cheapest first, making an extra unit of abatement cheaper if the current level of abatement is low. This is consistent with an increasing cost function ($\frac{c_i}{2}q_i^2$ from eq. (3.1) is increasing when $c_i, q_i > 0$). The cost parameter (c_i) is also a purely monetary measure.

4.1 Regional parameters

Collecting all the data necessary to produce realistic estimations of the parameters is a time-consuming process. Luckily, others have developed and used methods to calculate these parameters. One such method is presented by Osmani and Tol (2009). They use version 2.8 of the integrated assessment model Climate Framework for Uncertainty, Negotiation and Distribution (FUND) to estimate the parameters in a linear benefit/quadratic cost structure (Osmani & Tol, 2009) similar to the one I use in this thesis.

FUND is a huge and complicated model that takes everything from economic and demographical factors to a vide range of other consequences of global warming into consideration (Tol, 2006). A detailed description of the FUND model estimations are beyond the scope of this thesis, and I will therefore take the parameters presented by Osmani and Tol as given in the following chapters.

The payoff function I use is slightly different from the payoff function presented by Osmani and Tol (2009). I therefore need to revise the parameter estimations presented in their paper. This is inspired by Holtsmark (2013).

Osmani and Tol's payoff function with $n=16$ heterogeneous regions with subscript i, is:

$$\pi_i = \beta_i \sum_j^n R_j E_j - \alpha_i R_i^2 Y_i$$

Here β_i is the marginal damage costs of carbon dioxide emissions (in dollars per tonne of carbon), α_i is the abatement cost parameter (unitless), R_i is the relative emission reduction, E_i is unabated carbon dioxide emissions (in billion metric tonnes of carbon) and Y_i is gross domestic product (in billions US dollars). The 16 regions and the corresponding parameter estimations are presented in table 4.1, together with data on E_i and Y_i.

	α_i	β_i	E_i	Y_i
USA	0.01515466	2.19648488	1.647	10399
Canada	0.01516751	0.093156	0.124	807
Western Europe	0.01568	3.15719404	0.762	12575
Japan and South Korea	0.0156278	-1.42089104	0.525	8528
Australia and New Zealand	0.0151065	-0.05143806	0.079	446
Central and Eastern Europe	0.01465218	0.10131831	0.177	407
Former Soviet Union	0.01381774	1.27242378	0.811	629
Middle East	0.01434659	0.04737632	0.424	614
Central America	0.01486421	0.06652486	0.115	388
South America	0.015137	0.26839935	0.223	1351
South Asia	0.01436564	0.35566631	0.559	831
Southeast Asia	0.01484894	0.73159104	0.334	1094
China plus	0.01444354	4.35686225	1.431	2376
North Africa	0.01459959	0.96627119	0.101	213
Sub-Saharan Africa	0.01459184	1.07375825	0.145	302
Small Island States	0.01434621	0.05549814	0.038	55

Table 4.1 – The FUND model

Source: Osmani & Tol (2009, p. 461)

To avoid confusing the regions in Osmani and Tol's paper with my 16 individual countries, I will from now on index regions (r) with superscript and countries (i) with subscript.

Abatement (q^r) can be defined as emissions with no abatement (E^r) minus actual emissions (which I call Ω^r). Ω^r can also be written as $(1 - R^r)E^r$, where R^r is relative emission reduction, $0 \leq R^r \leq 1$ and E^r is unabated emissions. We can then define q^r (the variable in my payoff function) in terms of R^r and E^r (variables in Osmani and Tol's payoff function) as $q^r = R^rE^r$ (Holtsmark, 2013).[20]

b^r is the benefit parameter for region r in the payoff function I use in this thesis. b^r can be derived from β^r; the benefit parameter in Osmani and Tol's payoff function:

$$b^r \sum_{j=1}^{n} q^j = \beta^r \sum_{j}^{n} R^j E^j \quad \Longrightarrow \quad b^r = \beta^r$$

β^r is, according to Osmani and Tol (2009), the marginal damage costs of carbon dioxide emissions (in dollars per tonne of carbon), so b^r should then also be interpreted as the marginal damage costs of carbon dioxide emissions for region r.

Region *r*	**Population**	**CO$_2$ emissions**	**b^r**	**c^r**
USA	295 516 599	5 595 358	2.1965	116.19
Canada	32 312 000	562 796	0.0932	1592.12
Western Europe	402 671 727	3 390 758	3.1572	679.16
Japan and South Korea	175 911 000	1 701 107	-1.4209	967.07
Australia and New Zealand	24 528 700	400 763	-0.0514	2159.11
Central and Eastern Europe	108 652 994	744 885	0.1013	380.70
Former Soviet Union	279 020 948	2 431 749	1.2724	26.43
Middle East	258 924 895	1 643 744	0.0474	98.00
Central America	145 394 537	476 149	0.0665	872.18
South America	314 701 356	793 495	0.2684	822.46
South Asia	1 516 964 984	1 601 676	0.3557	76.41
Southeast Asia	565 208 630	1 059 961	0.7316	291.24
China plus	1 337 307 870	5 924 688	4.3569	33.52
North Africa	153 282 846	399 487	0.9663	609.69
Sub-Saharan Africa	740 545 039	640 478	1.0738	419.19
Small Island States	44 329 527	106 526	0.0555	1092.86

Table 4.2 – Region *r*'s total population and emissions, and the recalculated region parameters.

Total population in region *r* is the sum of the population in each individual country within region *r*.
Total emissions in region *r* are the sum of the emissions in each individual country within region *r*.
See table (4.3) below for a list of the countries belonging to each region.

Source: Population data: World Bank (2005b), emission data: World Bank (2005a), benefit parameter b_i and cost parameter c_i: Holtsmark (2013) based on Osmani &Tol (2009)

---

[20] $q^r = E^r - \Omega^r = E^r - (1 - R^r)E^r = R^rE^r$

The cost parameter for region r (c^r) corresponding to the payoff function in this thesis (eq. (3.9)) can also be calculated from Osmani and Tol's payoff function (where α^r is the cost parameter for region r). Keeping in mind that $q^r = R^r E^r$, q^r is abatement in eq. (3.9), R^r is defined by Osmani and Tol as relative emission reduction, E^r as unabated carbon dioxide emissions and Y^r as gross domestic product (in billions US dollars), we can calculate c^r (Holtsmark, 2013):

$$\frac{c^r}{2}(q^r)^2 = \alpha^r (R^r)^2 Y^r \quad \Longrightarrow \quad c^r = 2\alpha^r \frac{Y^r}{(E^r)^2}$$

The results of these calculations are presented in table 4.2 above, together with population and emissions data for the 16 regions.

4.2 The parameters of country *i*

Holtsmark (2013, p. 338) shows that aggregation of countries is only appropriate as long as the aggregated units actually act as a union. Otherwise such aggregations can give misleading results. The relevant parties to an international climate treaty are independent states, not regions consisting of several states (see section 2.3).

Osmani and Tol estimated the payoff functions of $n = 16$ world regions (r) (2009). I want to find the parameters corresponding to $N = 16$ independent countries (i). I have chosen to use the 16 countries with the highest emission of CO_2 in 2005 (World Bank, 2005a).

I make the simplifying approximation that all individuals within region r experience the same consequences of global warming because of similarities in natural resources, climate and other geographical aspects within the region. If there are cultural similarities, local preferences are similar within region r, and the individuals within the region have similar incomes, they may also value the damages avoided by abatement the same way. With these assumptions, it is reasonable to assume that every single individual within region r have the same, identical marginal benefit from abatement, which I label $\widetilde{b^r}$.

Finally I assume that p^r is the total population in region r and that the region benefit parameter is simply the sum of the benefit parameters of every individual within region r:

$$p^r \, \widetilde{b^r} = b^r \tag{4.1}$$

p_i^r is the total population in country *i* within region *r*. Since I have assumed that the region's benefit parameter is the sum of the region population's individual benefit parameters, country i's benefit parameter is the sum of the individual benefit parameters of country i's population: [21]

$$p_i^r\, \widetilde{b^r} = b_i \qquad \Longrightarrow \qquad p_i^r \frac{b^r}{p^r} = b_i \tag{4.2}$$

This formula is later used for the calibration of country-specific benefit parameters.

Let's now assume that each country *i* within region *r* has their own individual marginal abatement cost function and that the region's marginal abatement cost function is an aggregation of these functions.[22] The marginal cost in region *r* is not the sum of the cost of one unit of abatement in every country, but the cost of abating one unit in *one* country. Thus, the aggregated marginal cost is *not* the sum of the individual marginal costs at given levels of abatement.

It makes more sense to think of the aggregated marginal abatement cost as the sum of individual abatement decisions at a *given* marginal cost of abatement. The marginal cost of abatement corresponds to a specific level of abatement for each country *i*. With heterogeneous cost functions, a given marginal cost of abatement corresponds to a different level of abatement for each country. At this given marginal cost, aggregated abatement is the sum of these different abatement levels. I therefore use a technique called horizontal summation to separate country *i*'s abatement cost parameter from region *r*'s. This technique is explained very simply by Robert Hansen (2007).

The marginal cost of abatement for country *i* is the derivative of country *i*'s payoff function (see eq. (3.9)) with respect to abatement (q_i).

$$MAC_i = c_i q_i \quad \Longleftrightarrow \quad q_i = \frac{1}{c_i} MAC_i$$

Similarly, the marginal cost of abatement for region *r* is the derivative of region *r*'s payoff function (see eq. (3.9)) with respect to abatement (q^r).

---

[21] Where all $\widetilde{b_r}$ are identical within region *r* and equal to $\frac{b^r}{p^r}$ (from eq. (4.1))

[22] The marginal cost of abatement (MAC) for country *i* is $c_i q_i$

$$MAC^r = c^r q^r \iff q^r = \frac{1}{c^r} MAC^r$$

Since we use horizontal summation, we know that aggregated abatement (i.e. abatement in region *r*) is the sum of individual abatement decisions (i.e. abatement in country *i*) at a *given* marginal cost of abatement; where $MAC_i = MAC^r = MAC$.

$$q^r = \sum_{all\ i \in r} q_i \implies \frac{1}{c^r} MAC = \sum_{all\ i \in r} \left(\frac{1}{c_i}\right) MAC \implies$$

$$\frac{1}{c^r} = \sum_{all\ i \in r} \left(\frac{1}{c_i}\right) \tag{4.3}$$

I further assume that production technologies are similar for all countries *i* within region *r*. Consider countries *v* and *w*; they are both part of region *r* and total emissions in country *v* are higher than total emissions in country *w*. Production technologies are similar; there are differences in total emissions either because country *w* abates more than country *v* (i.e. country *v* is less environmentally friendly) or because in country *v* the polluting activity is of a larger magnitude to begin with than in country *w* (i.e. country *v* has a larger polluting sector).

Regardless of why emissions are higher in country *v* than in country *w*, one additional unit of abatement should be cheaper for country *v* than for country *w*. If the economic conditions within region *r* are similar, it is not unreasonable to assume that country *v* will find cheaper ways to reduce emission than country *w*, either because country *v* has lower total abatement than country *w* (since the marginal cost of abatement is increasing with abatement, every additional unit of abatement is more expensive for country *v* than for country *w*)[23] or because country *w* has a smaller polluting sector than country *v*, and might be harder pressed to reduce emissions simply because it has lower emissions.

In terms of the model, this means that country *w*'s marginal cost of abatement should be *higher* than country *v*'s marginal cost of abatement. The marginal abatement cost for country *i* depends on country *i*'s cost parameter (c_i) and abatement in country *i* (q_i); a higher marginal cost of abatement could either come from a higher cost parameter or a higher level of abatement.

---

[23] The marginal cost of abatement ($MAC_i = c_i q_i$) is increasing in abatement because the derivative of the marginal cost is positive: $MAC_i' = c_i > 0$ if $c_i > 0$

If the difference between total emissions in country *v* and *w* stem only from differences in abatement and the two countries have equally large polluting activities, then the marginal cost in country *w* is higher than the marginal cost in country *v* simply because q_v is lower than q_w. But if the countries have identical technology, similar economies and the same magnitude of polluting activities, their cost of abatement should be identical *for every given level of abatement*, implying that $c_v = c_w$.

If country *w*'s polluting activity is smaller than country *v*'s, country *w* should have a higher marginal cost of abatement than country *v for every given level of abatement*. Since country *i*'s marginal cost of abatement depends only on country *i*'s cost parameter (c_i) and abatement in country *i* (q_i), this implies that country *w* should have a higher cost parameter than country *v*.

This means that the relative size of country *i*'s polluting activity is what determines the value of country *i*'s cost parameter relative to the cost parameter in region *r*, given that production technologies and economic conditions are similar within region *r*.

I have considered two methods to determine the relative size of country *i*'s polluting activity when country *i* is a part of region *r*. Both methods are very simple.

I first define x_i as country *i*'s proportion of region *r*'s total emissions.[24] I then assume that x_i can be interpreted as the size of country *i*'s polluting activity relative to the total polluting activity in region *r*. The problem with this interpretation is that country *i* can have relatively low emissions and still have a relatively high level of polluting activities – because it can have a *high level of abatement* already. This method will thus tend to give too much weight to countries that are *lax* (relative to the others within the same region) when it comes to abatement, which gives these countries a lower cost parameter than they should have.

The second approach is to define y_i as country *i*'s proportion of region *r*'s GDP,[25] and then interpret y_i as the size of country *i*'s polluting activity relative to the total polluting activity in region *r*. The problem with this interpretation is that country *i* can have a relatively high GDP and still have a relatively low level of polluting activities – because relatively large parts of *i*'s

---

[24] $x_i = \frac{Emissions\ in\ country\ i}{Emissions\ in\ region\ r}$ so that $\sum_{all\ i \in r} x_i = 1$ – data on emissions is found at the World Bank (2005a)

[25] $y_i = \frac{Country\ i's\ GDP}{Region\ r's\ GDP}$ so that $\sum_{all\ i \in r} y_i = 1$ – data on GDP is found at the World Bank (2005c)

national income may come from *non-polluting activities*. This method will thus tend to give too little weight to countries that get most of their income from polluting activities (relative to those that are more ‘environmentally friendly’), which gives these countries a higher cost parameter than they should have.

If I assume that all countries behave according to the model, we know from the discussion in previous sections that countries abate very little when they make decisions unilaterally.[26] Thus, there should be no large differences in existing abatement level and therefore no large errors when using relative emissions as the measure of the relative size of country *i*’s polluting activity.

It is not possible to tell from the model whether countries within region *r* should have *equally polluting* income-generating production or not. The errors from using GDP as the measure of the relative size of country *i*’s polluting activity could therefore be substantial.

I therefore conclude that x_i is my preferred measure for determining the relative size of country *i*’s polluting activity when country *i* is a part of region *r*.

We know that $\frac{1}{c^r} = \sum_{all\ i \in r} \left(\frac{1}{c_i}\right)$ and that x_i is the relative size of country *i*’s polluting activity. The inverse of country *i*’s abatement cost parameter $(\frac{1}{c_i})$ is the portion x_i of the inverse of region *r*’s cost parameter $(\frac{1}{c^r})$.[27] Country *i*’s abatement cost parameter can then be expressed in terms of region *r*’s cost parameter (c^r) and x_i:

$$c_i = \frac{1}{x_i} c^r \qquad (4.4)$$

4.3 The value of the parameters b_i and c_i

I have presented the theoretical and methodological foundation for my calculations of the specific parameters for each country *i*. I will now put the method to use and calculate the values of the parameters for all 16 countries.

---

[26] The exact level of abatement for country *i* will of course depend on country *i*’s cost and benefit parameters.

[27] $\frac{1}{c_r} MAC = \sum_{all\ i \in r} \left(\frac{1}{c_i}\right) MAC = \sum_{all\ i \in r} \left(x_i \frac{1}{c_r}\right) MAC = \frac{1}{c_r} \sum_{all\ i \in r} (x_i)\, MAC = \frac{1}{c_r} MAC$

Table 4.3 presents all the countries belonging to each region. The 16 countries (including the EU) with the highest total emission of CO_2 in 2005 (the World Bank, 2005a) are chosen as the parties in the international climate treatise model. The EU is treated as a single player in this model because, as I mentioned in chapter 1, the European Union can make decisions that are binding for its member states and they negotiate internationally as a single unit. While my data is from 2005 and Bulgaria and Romania did not become members of the union until 2007, they are still included as members of the union. The data on population and emission from 2005 is only used to derive each country *i*'s individual parameters from region *r*'s parameters. I have found the cost and benefit parameters for each member of the EU based on data from 2005. The parameters for the European Union are then found by aggregating its members' parameters.

REGION	COUNTRY
USA	**United States of America**
Canada	**Canada**
Western Europe	Andorra, **Austria***, **Belgium***, **Cyprus***, **Denmark***, **Finland***, **France***, **Germany***, **Greece***, Iceland, **Ireland***, **Italy***, Liechtenstein, **Luxembourg***, **Malta***, Monaco, **Netherlands***, Norway, **Portugal***, San Marino, **Spain***, **Sweden***, Switzerland, **United Kingdom***
Japan and South Korea	**Japan**, **South Korea**
Australia and New Zealand	**Australia**, New Zealand
Central and Eastern Europe	Albania, Bosnia and Herzegovina, **Bulgaria***, Croatia, **Czech Republic***, **Hungary***, FYR Macedonia, **Poland***, **Romania***, **Slovak Republic***, **Slovenia***, Yugoslavia
Former Soviet Union	Armenia, Azerbaijan, Belarus, **Estonia***, Georgia, Kazakhstan, **Latvia***, **Lithuania***, Moldova, **Russia**, Tajikistan, Turkmenistan, **Ukraine**, Uzbekistan
Middle East	Bahrain, **Iran**, Iraq, Israel, Jordan, Kuwait, Lebanon, Oman, Qatar, **Saudi Arabia**, Syria, Turkey, United Arab Emirates, West Bank and Gaza, Yemen
Central America	Belize, Costa Rica, El Salvador, Guatemala, Honduras, **Mexico**, Nicaragua, Panama
South America	Argentina, Bolivia, **Brazil**, Chile, French Guiana, Guyana, Paraguay, Peru, Suriname, Uruguay, Venezuela
South Asia	Afghanistan, Bangladesh, Bhutan, **India**, Nepal, Pakistan, Sri Lanka
Southeast Asia	Brunei, Cambodia, East Timor, **Indonesia**, Laos, Malaysia, Myanmar, Papua New Guinea, Philippines, Singapore, Taiwan, Thailand, Vietnam
China plus	**China**, Hong Kong, North Korea, Macau, Mongolia
North Africa	Algeria, Egypt, Libya, Morocco, Tunisia, Western Sahara
Sub-Saharan Africa	Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Congo-Brazzaville, Congo-Kinshasa, Cote d'Ivoire, Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea- Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mauritania, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, Somalia, **South Africa**, Sudan, Swaziland, Tanzania, Togo, Uganda, Zambia, Zimbabwe
Small Island States	Antigua and Barbuda, Aruba, Bahamas, Barbados, Bermuda, Comoros, Cuba, Dominica, Dominican Republic, Fiji, French Polynesia, Grenada, Guadeloupe, Haiti, Jamaica, Kiribati, Maldives, Marshall Islands, Martinique, Mauritius, Micronesia, Nauru, Netherlands Antilles, New Caledonia, Palau, Puerto Rico, Reunion, Samoa, Sao Tome and Principe, Seychelles, Solomon Islands, St Kitts and Nevis, St Lucia, St Vincent and Grenadines, Tonga, Trinidad and Tobago, Tuvalu, Vanuatu, Virgin Islands

Table 4.3 – List of countries within the 16 regions in FUND

Countries marked in **bold** are the 16 individual countries, including members of the European Union
* Members of the European Union
Source: (TABLES, 2006)

These 16 countries belong to 14 of the 16 world regions in the FUND model. Most countries within Western Europe and some in both Central and Eastern Europe and the Former Soviet Union are countries that belong to the European Union. Both the US and Canada are the only countries in their respective region, while Japan, South Korea, Australia and China are part of quite uniform regions with few other countries. Only South Africa belongs to a region with a very large number of other countries. The two regions to which none of the relevant countries belong are North Africa and the Small Island States.

Looking at the regions in table 4.3, there are obviously differences between some countries within the same region. South Africa is unique when it comes to income, culture, industry, etc., among the South African countries. Latvia, Lithuania and Estonia differ from the other Former Soviet Union countries and Iran and Saudi Arabia have both economic and cultural differences.

Still, keeping in mind that the calculations in this thesis are simply an interesting first approach to the problem of deriving the benefit and cost parameters for individual countries, it does not seem unreasonable to assume that the people within each of the 14 relevant regions have equal individual benefit parameters. Within each region the conditions, both economically, culturally and geographically, are sufficiently similar for me to calculate country *i*'s benefit parameter (b_i) using eq. (4.2) in this thesis.

So, to find the benefit parameters (b_i) for each of the individual countries, I have to use the region parameters b^r and the population size for the 16 regions (*r*) and the 16 independent countries (*i*). The population data for the regions are found in table 4.2 above, while the data for the individual countries will be presented in table 4.4.

I have argued that relative emissions can be used to determine the size of a country's polluting activity relative to that of the region's. While real countries do not behave according to any theoretical model, I still think that differences in abatement level within each region are small enough to justify using relative emissions as an approximation for relative size of polluting activities – at least for the purposes of this thesis. I will therefore use equation (4.4) to calculate country *i*'s abatement cost parameter.

To find the cost parameters, I have to use the region cost parameters c^r, total emission in the regions (*r*) and emissions in the 16 independent countries (*i*).[28] The results are presented in table 4.4 below.

---

[28] Total emission in region *r* is found by adding the emissions in each country within the region.

As we can see from this table, China, the US and the EU are by far the countries that have the highest marginal willingness to pay for abatement. They also have relatively low cost parameters. Japan, the Republic of Korea and Australia on the other hand, actually have a negative marginal willingness to pay for abatement (i.e. they are willing to pay for higher emissions). The benefit parameters are negative because the benefit parameters for their respective regions are negative in the FUND model (see table 4.1). These three countries also have high cost parameters.

Country *i*	**Region *r***		**Population**	**CO_2 emissions**	**b_i**	**c_i**	**x_i**
China	*China plus*		1 303 720 000	5 790 017	4.2475	34.30	0.9773
United States of America	*USA*		295 516 599	5 595 358	2.1965	116.19	1
EU			492 447 439	4 023 176	3.1799	225.48	-
	Western Europe	Austria	8 227 829	74 385	0.0645	30 958.71	0.0219
		Belgium	10 478 617	107 117	0.0822	21 498.67	0.0316
		Cyprus	1 032 562	7 503	0.0081	306 939.16	0.0022
		Denmark	5 419 432	46 868	0.0425	49 135.24	0.0138
		Finland	5 246 096	54 605	0.0411	42 172.96	0.0161
		France	63 001 253	391 826	0.4940	5 877.26	0.1156
		Germany	82 469 422	809 597	0.6466	2 844.46	0.2388
		Greece	11 103 965	98 675	0.0871	23 337.82	0.0291
		Ireland	4 159 914	43 212	0.0326	53 292.39	0.0127
		Italy	58 607 043	471 400	0.4595	4 885.16	0.1390
		Luxembourg	465 158	11 327	0.0036	203 301.24	0.0033
		Malta	403 837	2 699	0.0032	853 257.50	0.0008
		Netherlands	16 319 868	172 228	0.1280	13 371.04	0.0508
		Portugal	10 549 424	65 309	0.0827	35 260.95	0.0193
		Spain	43 398 143	353 462	0.3403	6 515.17	0.1042
		Sweden	9 029 572	51 551	0.0708	44 671.90	0.0152
		United Kingdom	60 224 307	542 474	0.4722	4 245.12	0.1600
	Central and Eastern Europe	Bulgaria	7 739 900	47 909	0.0072	5 919.05	0.0643
		Czech Republic	10 235 828	120 736	0.0095	2 348.74	0.1621
		Hungary	10 087 065	57 880	0.0094	4 899.41	0.0777
		Poland	38 165 445	303 767	0.0356	933.54	0.4078
		Romania	21 634 371	94 961	0.0202	2 986.27	0.1275
		Slovak Republic	5 387 001	39 175	0.0050	7 238.82	0.0526
		Slovenia	2 000 474	15 871	0.0019	17 867.92	0.0213
	Former Soviet Union	Estonia	1 346 097	17 462	0.0061	3 680.57	0.0072
		Latvia	2 300 512	7 176	0.0105	8 956.00	0.0030
		Lithuania	3 414 304	14 001	0.0156	4 590.60	0.0058
Russian Federation	*Former Soviet Union*		143 150 000	1 615 684	0.6528	39.78	0.6644
India	*South Asia*		1 140 042 863	1 411 128	0.2673	86.73	0.8810
Japan	*Japan and South Korea*		127 773 000	1 238 188	-1.0321	1 328.63	0.7279
Canada	*Canada*		32 312 000	562 796	0.0932	1 592.12	1
Korea, Rep.	*Japan and South Korea*		48 138 000	462 918	-0.3888	3 553.73	0.2721
Iran, Islamic Rep.	*Middle East*		69 732 007	458 866	0.0128	351.05	0.2792
Mexico	*Central America*		106 483 757	432 666	0.0487	959.83	0.9087
South Africa	*Sub-Saharan Africa*		47 198 469	408 199	0.0684	657.72	0.6373
Australia	*Australia and New Zealand*		20 394 800	367 393	-0.0427	2 355.22	0.9167
Saudi Arabia	*Middle East*		24 041 116	367 034	0.0044	438.89	0.2233
Brazil	*South America*		185 986 964	349 967	0.1586	1 864.80	0.4410
Ukraine	*Former Soviet Union*		47 105 150	339 029	0.2148	189.57	0.1394
Indonesia	*Southeast Asia*		227 303 175	336 312	0.2942	917.91	0.3173

Table 4.4 – Population, emissions, relative emissions and parameter values for country *i*.

CO_2 emissions measured in (kt), b_i is a benefit parameter, c_i is a cost parameter and x_i is emissions in country *i* relative to emissions in region *r* (see table (4.2)*)*.

The population, emission and parameters for the EU is found by first calculating the values for all the individual member countries (relative to the region they belong to), then adding them together to find the aggregated values for the EU as a whole. The cost parameter for EU was aggregated using horizontal summation (see sec. 4.2).

Source: Population data (World Bank, 2005b), emission data (World Bank, 2005a), the benefit parameters b_i and cost parameters c_i are calculated from region parameters found in (Osmani & Tol, 2009) and (Holtsmark, 2013)

5 Numerical results

When I solve the international negotiation model for 16 heterogeneous countries applying my own parameters (see table 4.4) to Holtsmark's spreadsheets (2013), I get 65 519 possible coalitions with at least two members. In addition, I have to consider the privately optimal outcome where no countries cooperate.

In the following I will consider a world consisting only of the following 16 countries; China, the United States of America, the European Union, the Russian Federation, India, Japan, Canada, the Republic of Korea, the Islamic Republic of Iran, Mexico, South Africa, Australia, Saudi Arabia, Brazil, Ukraine and Indonesia. When I use terms such as *global* and *socially optimal* I refer to this rather limited world. While this model in no way represents the real world, these results might give some useful insights none the less.

A coalition is internally stable if no single member profits from withdrawing from the coalition (see discussion of internal stability in section 2.5 and chapter 3). In other words, equation (3.11) must hold for each of the 16 individual countries. Of the 65 519 possible coalitions only 517 are internally stable. The non-cooperative outcome is also internally stable according to Holtsmark's spreadsheets; making a total of 518 internally stable outcomes. This is only 0.8% of all possible outcomes.

Internal stability is in fact a subcategory of potential internal stability; all coalitions that are internally stable are also potentially internally stable. To be potentially internally stable, the sum of the net gain of membership for all members must be positive (see discussion of potential internal stability in section 2.5 and chapter 3). In other words, equation (3.13) must hold. This condition ensures that the countries that gain from being part of the coalition without side payments can make side payments to countries that do not gain from participation without side payments and entice them to be part of the coalition. There are in total 53 796 potentially internally stable outcomes; 99% of them are not internally stable without side payments.

The overwhelming majority – 82% – of all possible coalitions are potentially internally stable. This leaves us with a total of 11 724 coalitions that are not potentially internally stable.

When there is no cooperation, each country abates at the privately optimal level of abatement, called the BaU level of abatement or the non-cooperative abatement level. This level, q_i^*, is found by solving eq. (3.2) for each of the 16 countries.

There is one coalition that includes all 16 countries as members. This grand coalition gives the socially optimal level of abatement, q_i^{**}, which is found by solving eq. (3.5) for each of the 16 countries.

		BaU		Social optimum	
		Payoff	q_i^*	Payoff	q_i^{**}
	GLOBAL	***1.45539649***	***0.17722837***	***4.75374916***	***0.95308205***
1	China	*0.48978122*	*0.12383428*	*2.59757376*	*0.29083478*
2	United States	*0.36852038*	*0.01890438*	*1.66521857*	*0.08585532*
3	European Union	*0.54114871*	*0.01410297*	*2.81005712*	*0.04424158*
4	Russian Federation	*0.11033784*	*0.01641038*	*-0.62861509*	*0.2507705*
5	India	*0.0469645*	*0.00308227*	*-0.31892132*	*0.11502107*
6	Japan	*-0.18331313*	*-0.00077680*	*-1.02109733*	*0.00750815*
7	Canada	*0.01651496*	*0.00005854*	*0.05757609*	*0.00626556*
8	Korea, Rep.	*-0.06893279*	*-0.00010941*	*-0.3845868*	*0.00280706*
9	Iran, Islamic Rep.	*0.00226217*	*0.00003636*	*-0.12956549*	*0.02841594*
10	Mexico	*0.00863034*	*0.00005074*	*-0.00541958*	*0.01039296*
11	South Africa	*0.01212567*	*0.00010405*	*-0.01042087*	*0.01516677*
12	Australia	*-0.00757467*	*-0.00001815*	*-0.06185793*	*0.0042355*
13	Saudi Arabia	*0.00077997*	*0.00001003*	*-0.10917268*	*0.02272907*
14	Brazil	*0.02810577*	*0.00008506*	*0.12449929*	*0.0053494*
15	Ukraine	*0.03794879*	*0.00113312*	*-0.0577278*	*0.05262071*
16	Indonesia	*0.05209677*	*0.00032053*	*0.22620922*	*0.01086768*

Table 5.1 – BaU and social optimum

Abatement is measured in billion tonnes CO_2 and payoff in US dollars

Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

As we can see from table 5.1; while the socially optimal situation is superior when it comes to global payoff, only 6 of the 16 countries actually prefer this outcome to the BaU situation. The remaining 10 countries have a lower individual payoff in the socially optimal situation than in the non-cooperative case.

The socially optimal situation is neither internally nor potentially internally stable. The non-cooperative outcome, however, is defined as internally stable in Holtsmark's spreadsheets (2013). This is because, even though there is no cooperation and thus no coalition, the countries do not gain from leaving their own private 'coalition' of one.

The non-cooperative and the socially optimal outcomes are only two out of 65 520 different outcomes. How do the other outcomes compare to these two? In the following sections, I will use two indexes to measure the relative performance of the different coalitions (Holtsmark, 2013). These will be presented in section 5.1.

5.1 The efficiency index and the environmental index

The efficiency index is defined in equation (5.1) and measures the relative performance of each coalition in terms of global payoff.

$$I_{eff}(K) = \frac{\theta^{K} - \theta^{BaU}}{\theta^{SO} - \theta^{BaU}} \qquad (5.1)$$

$\theta(K)$ is global payoff when there is a coalition K, $\theta(BaU)$ is global payoff in the non-cooperative outcome and $\theta(SO)$ is global payoff in the socially optimal outcome. The index is equal to zero at the value of the global payoff in the non-cooperative outcome, i.e. when there is no cooperation. It equals one at the value of the global payoff in socially optimal outcome, i.e. when there is one grand coalition including all 16 countries as members.

The environmental index measures the performance of each coalition in terms of global abatement and is defined in equation (5.2).

$$I_{env}(K) = \frac{\varepsilon^{K} - \varepsilon^{BaU}}{\varepsilon^{SO} - \varepsilon^{BaU}} \qquad (5.2)$$

$\varepsilon(K)$ is global abatement when there is a coalition K, $\varepsilon(BaU)$ is global abatement in the non-cooperative outcome and $\varepsilon(SO)$ is global abatement in the socially optimal outcome. The index equals zero at the level of abatement achieved in non-cooperative outcome and is equal to one at the level of abatement achieved in the socially optimal outcome.

Figure 5.1 below depicts the efficiency index for each of the 65 520 possible outcomes. The outcomes are sorted by the number of members in each coalition (from zero to 16). The internally stable coalitions are marked by grey circles, the potentially internally stable coalitions are marked by black triangles and all other coalitions are marked by light grey squares.

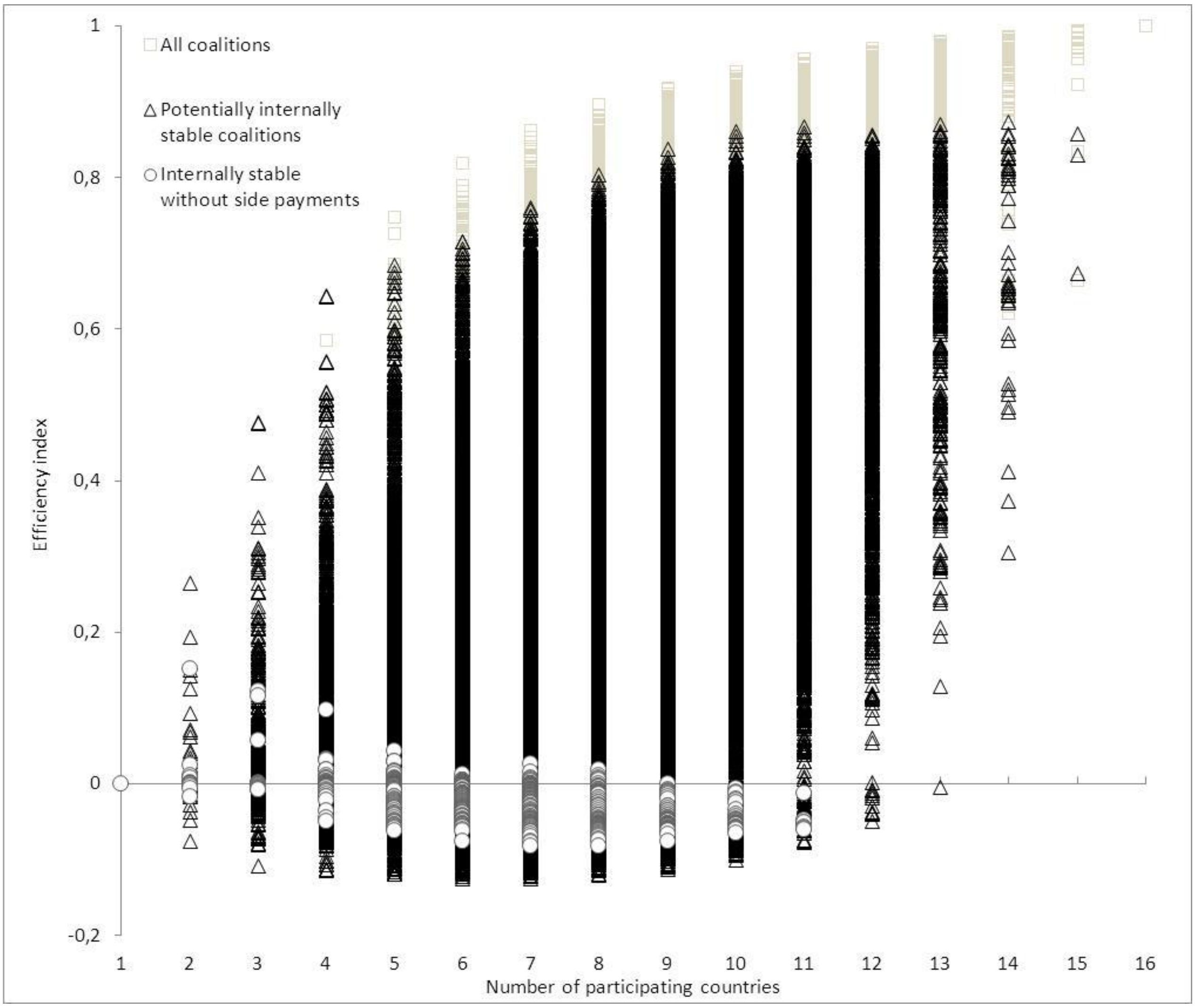

Figure 5.1 – The efficiency index depending on coalition size
The relative performance of all possible coalitions according to the efficiency index; eq. (5.1)
Circles are internally stable coalitions, triangles are potentially internally stable coalitions and squares are the remaining coalitions
NB: Some markers (circles, triangles or squares) are spaced so close together that they may appear to be a solid column
Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

Figure 5.2 below depicts the environmental index for each of the 65 520 possible outcomes. The outcomes are also here sorted by the number of members in a coalition. Internally stable coalitions are marked by grey circles and potentially internally stable coalitions are marked by black triangles. All other coalitions are marked by light grey squares.

Some of the results from these figures will be presented in the next three sections and then discussed in section 5.5.

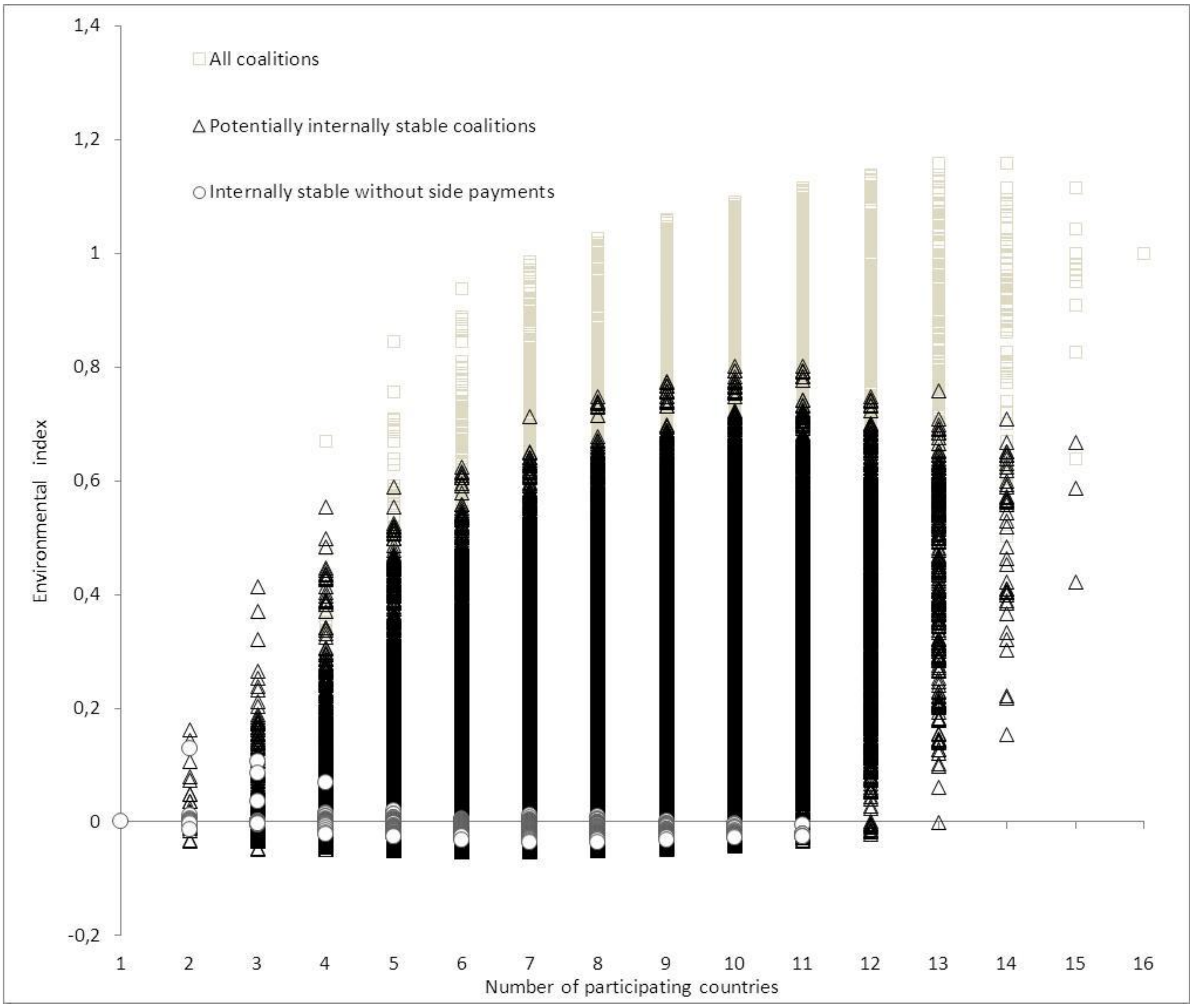

Figure 5.2 – The environmental index depending on coalition size
The relative performance of all possible coalitions according to the environmental index; eq. (5.2)
Circles are internally stable coalitions, triangles are potentially internally stable coalitions and squares are the remaining coalitions
NB: Some markers (circles, triangles or squares) are spaced so close together that they may appear to be a solid column
Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

5.2 Internally stable coalitions

As we can see from figure 5.1, internally stable coalitions generally perform poorly with regards to global payoff. In addition, the overall trend is that the larger the coalition is the lower is the efficiency index. Larger coalitions actually perform worse than the small ones.

Likewise, we can see from figure 5.2 that internally stable coalitions also perform relatively poorly with regards to abatement. The overall trend is that larger coalitions have lower environmental indexes. This means that larger coalitions abate less than the small ones.

The internal stability condition (the left-hand side of equation (3.11)) for each member of the coalition is illustrated by the IS column in tables 5.2 through 5.9. The numbers represent each member's net gain from being a member of the coalition. If the net gain is positive for all members, they all want to stay in the coalition and the coalition is internally stable. The net gain has been calculated for each member of each of the 65 519 possible coalitions using Holtsmark's spreadsheets.

The best performing internally stable coalition, both with regards to global payoff and global abatement, is a coalition with only two members; China and the US (table 5.2 below). It has an efficiency index of 0.1526; the coalition provides only 15% of the maximum achievable welfare gain of an agreement between these 16 countries. The environmental index is 0.1297.

		Payoff	*Abatement*	*IS*
	GLOBAL	1.958619	0.277823	-
1	**China**	**0.574723**	**0.187873**	**0.084941**
2	**United States**	**0.431545**	**0.055461**	**0.063025**
3	European Union	0.861032	0.014103	-
4	Russian Federation	0.176006	0.016410	-
5	India	0.073855	0.003082	-
6	Japan	-0.287134	-0.000777	-
7	Canada	0.025890	0.000059	-
8	Korea, Rep.	-0.108047	-0.000109	-
9	Iran, Islamic Rep.	0.003546	0.000036	-
10	Mexico	0.013530	0.000051	-
11	South Africa	0.019010	0.000104	-
12	Australia	-0.011874	-0.000018	-
13	Saudi Arabia	0.001223	0.000010	-
14	Brazil	0.044062	0.000085	-
15	Ukraine	0.059558	0.001133	-
16	Indonesia	0.081694	0.000321	-

Table 5.2 – The overall best performing internally stable coalition (BIS): China and the US

Countries marked in **bold** are members of the coalition
A coalition is IS if all the numbers in the IS column are ≥ 0.
A coalition is PIS if the sum of all numbers in the IS column is ≥ 0
IS: Internal stability. PIS: Potential internal stability
Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

As we can see, the numbers in the IS column are positive for both members, confirming that this coalition is in fact internally stable.

Looking at figure 5.1 and 5.2, we see that there are several internally stable coalitions with a large number of members. The largest coalitions have as many as 11 members. One of these is presented in table 5.3 below; Russia, Japan, the Republic of Korea, Iran, Mexico, South Africa, Australia, Saudi Arabia, Brazil, Ukraine and Indonesia are members of this coalition. This large coalition (LIS) has an environmental index of -0,0227 and an efficiency index of -0,0515.

		Payoff	Abatement	IS
	GLOBAL	1.285669	0.159617	-
1	China	0.414976	0.018904	-
2	United States	0.329836	0.018904	-
3	European Union	0.485145	0.014103	-
4	**Russian Federation**	**0.104196**	**-0.000223**	**0.001336**
5	India	0.042257	0.003082	-
6	**Japan**	**-0.164736**	**-0.000007**	**0.042676**
7	Canada	0.014874	0.000059	-
8	**Korea, Rep.**	**-0.062064**	**-0.000002**	**0.006164**
9	**Iran, Islamic Rep.**	**0.002037**	**-0.000025**	**0.000006**
10	**Mexico**	**0.007774**	**-0.000009**	**0.000093**
11	**South Africa**	**0.010924**	**-0.000014**	**0.000181**
12	**Australia**	**-0.006822**	**-0.000004**	**0.000074**
13	**Saudi Arabia**	**0.000702**	**-0.000020**	**0.000001**
14	**Brazil**	**0.025319**	**-0.000005**	**0.001015**
15	**Ukraine**	**0.034287**	**-0.000047**	**0.001525**
16	**Indonesia**	**0.046962**	**-0.000010**	**0.003421**

Table 5.3 – An internally stable coalition with 11 members (LIS)
Countries marked in **bold** are members of the coalition
A coalition is IS if all the numbers in the IS column are ≥ 0
A coalition is PIS if the sum of all numbers in the IS column is ≥ 0
IS: Internal stability. PIS: Potential internal stability
Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

Looking at the IS column confirms that this coalition is in fact internally stable. The coalitions efficiency index and environmental index are negative; it has a lower global payoff and a lower global abatement than the non-cooperative solution (see table 5.1). In fact, taking a closer look at the abatement column, we see that the actual abatement level is negative for all members; the coalition members actually increase their emissions relative to the BaU level.

The worst performing internally stable coalition with regards to both global payoff and global abatement (WIS) is presented in table 5.4 below. Russia, India, Japan, the Republic of Korea, Mexico, South Africa and Ukraine are members of this coalition. This coalition has an efficiency index of -0.0817 and an environmental index of -0.0356; the coalition results in an even lower global payoff and lower global abatement than the coalition from table 5.3.

As many as 83% of all internally stable coalitions have negative efficiency indexes and 84% have negative environmental indexes (see figures 5.1 and 5.2). This implies that both in terms of global abatement and global payoff no cooperation would be preferable in some cases.

		Payoff	Abatement	IS
	GLOBAL	1.185909	0.149645	-
1	China	0.372623	0.123834	-
2	United States	0.307934	0.018904	-
3	European Union	0.453437	0.014103	-
4	**Russian Federation**	**0.097330**	**-0.004244**	**0.000209**
5	**India**	**0.039839**	**-0.001947**	**0.001334**
6	**Japan**	**-0.154455**	**-0.000127**	**0.047426**
7	Canada	0.013944	0.000059	-
8	**Korea, Rep.**	**-0.058190**	**-0.000048**	**0.006836**
9	Iran, Islamic Rep.	0.001910	0.000036	-
10	**Mexico**	**0.007273**	**-0.000176**	**0.000081**
11	**South Africa**	**0.010220**	**-0.000257**	**0.000163**
12	Australia	-0.006396	-0.000018	-
13	Saudi Arabia	0.000659	0.000010	-
14	Brazil	0.023730	0.000085	-
15	**Ukraine**	**0.032070**	**-0.000891**	**0.001470**
16	Indonesia	0.043981	0.000321	-

Table 5.4– The overall worst performing internally stable coalition (WIS)
Countries marked in **bold** are members of the coalition
A coalition is IS if all the numbers in the IS column are ≥ 0.
A coalition is PIS if the sum of all numbers in the IS column is ≥ 0
IS: Internal stability. PIS: Potential internal stability
Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

5.3 Potentially internally stable coalitions

Looking again at figure 5.1 we see that while the internally stable coalitions generally perform badly, there are many potentially internally stable coalitions that perform relatively well. A coalition is potentially internally stable if the potential internal stability condition (equation (3.13)) holds; the *sum* of the net gains of cooperation for all members of a coalition must be positive. Thus, the coalition is potentially internally stable if the sum of the net gains for each member (found in the IS column in tables 5.2 through 5.9) is positive.

The potentially internally stable coalition with the highest efficiency index (BeffPIS) – equal to 0.8735 – is a coalition consisting of 14 members (only the US and the Republic of Korea are not members). This coalition has an environmental index of 0.7081 and is depicted in table 5.5 below.

The sum of the net gains of cooperation for all members is positive – confirming that the coalition is potentially internally stable. The countries with positive net gains must pay the countries with negative net gains to stay in the agreement. In this case China, the EU and Japan have to make side payments to Russia, India, Canada, Iran, Mexico, South Africa, Australia, Saudi Arabia, Brazil, Ukraine and Indonesia in order to entice them to be part of the coalition.

		Payoff	Abatement	IS
	GLOBAL	4.336650	0.726573	-
1	**China**	**2.113581**	**0.238132**	**0.813297**
2	United States	1.575155	0.018904	-
3	**European Union**	**2.162504**	**0.036225**	**0.776219**
4	**Russian Federation**	**-0.364241**	**0.205328**	**-0.683650**
5	**India**	**-0.190390**	**0.094178**	**-0.354485**
6	**Japan**	**-0.774981**	**0.006148**	**0.059648**
7	**Canada**	**0.046765**	**0.005130**	**-0.019729**
8	Korea, Rep.	-0.282534	-0.000109	-
9	**Iran, Islamic Rep.**	**-0.085744**	**0.023267**	**-0.094709**
10	**Mexico**	**0.000634**	**0.008510**	**-0.034137**
11	**South Africa**	**-0.000990**	**0.012418**	**-0.049471**
12	**Australia**	**-0.045215**	**0.003468**	**-0.014154**
13	**Saudi Arabia**	**-0.072806**	**0.018610**	**-0.075920**
14	**Brazil**	**0.097363**	**0.004380**	**-0.015033**
15	**Ukraine**	**-0.019882**	**0.043085**	**-0.163069**
16	**Indonesia**	**0.177431**	**0.008898**	**-0.026362**

Table 5.5 – The potentially internally stable coalition with the highest efficiency index (BeffPIS)

Countries marked in **bold** are members of the coalition
A coalition is IS if all the numbers in the IS column are ≥ 0.
A coalition is PIS if the sum of all numbers in the IS column is ≥ 0
IS: Internal stability. PIS: Potential internal stability
Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

We can see from figure 5.2 that many potentially internally stable coalitions perform relatively well with regards to global abatement. This means that if we allow for side payments in this model, the countries are able to accomplish quite a lot in terms of global abatement through international negotiations.

The potentially internally stable coalition resulting in the highest level of global abatement (BenvPIS) is a coalition consisting of 11 members (including all countries except Canada, Mexico, Brazil, Ukraine and Indonesia). The coalition has an environmental index of 0.8012. The efficiency index for this coalition is 0.8670 and is presented in table 5.6 below.

As we can see from the IS column the coalition is not internally stable without side payments. The sum of the net gains for all members is, however, positive; the coalition is potentially internally stable if we allow for side payments between members. Here China, the US, the EU, Japan and the Republic of Korea have to make side payments to Russia, India, Iran, South Africa, Australia and Saudi Arabia in order to entice them to be part of the coalition.

		Payoff	Abatement	IS
	GLOBAL	4.315074	0.798826	-
1	**China**	**2.168267**	**0.267232**	**0.690418**
2	**United States**	**1.393079**	**0.078888**	**0.169053**
3	**European Union**	**2.353897**	**0.040651**	**0.755139**
4	**Russian Federation**	**-0.534538**	**0.230419**	**-0.884599**
5	**India**	**-0.270819**	**0.105687**	**-0.451129**
6	**Japan**	**-0.856062**	**0.006899**	**0.052699**
7	Canada	0.074448	0.000059	-
8	**Korea, Rep.**	**-0.322427**	**0.002579**	**0.000262**
9	**Iran, Islamic Rep.**	**-0.109464**	**0.026110**	**-0.119314**
10	Mexico	0.038904	0.000051	-
11	**South Africa**	**-0.009198**	**0.013936**	**-0.062518**
12	**Australia**	**-0.051976**	**0.003892**	**-0.017845**
13	**Saudi Arabia**	**-0.092198**	**0.020885**	**-0.095620**
14	Brazil	0.126705	0.000085	-
15	Ukraine	0.171474	0.001133	-
16	Indonesia	0.234982	0.000321	-

Table 5.6 – The potentially internally stable coalition with the highest environmental index (BenvPIS)

Countries marked in **bold** are members of the coalition

A coalition is IS if all the numbers in the IS column are ≥ 0.

A coalition is PIS if the sum of all numbers in the IS column is ≥ 0

IS: Internal stability. PIS: Potential internal stability

Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

		Payoff	Abatement	IS
	GLOBAL	1.042589	0.135826	-
1	China	0.313928	0.123834	-
2	United States	0.277581	0.018904	-
3	European Union	0.409494	0.014103	-
4	**Russian Federation**	**0.085185**	**-0.013232**	**-0.009756**
5	**India**	**0.034712**	**-0.006069**	**-0.001365**
6	**Japan**	**-0.140287**	**-0.000396**	**0.045176**
7	Canada	0.012656	0.000059	-
8	**Korea, Rep.**	**-0.052852**	**-0.000148**	**0.006494**
9	**Iran, Islamic Rep.**	**0.001339**	**-0.001499**	**-0.000407**
10	Mexico	0.006614	0.000051	-
11	South Africa	0.009292	0.000104	-
12	**Australia**	**-0.005864**	**-0.000223**	**0.000029**
13	**Saudi Arabia**	**0.000282**	**-0.001199**	**-0.000320**
14	Brazil	0.021538	0.000085	-
15	Ukraine	0.029055	0.001133	-
16	Indonesia	0.039916	0.000321	-

Table 5.7 – The overall worst performing coalition (WT)

Countries marked in **bold** are members of the coalition

A coalition is IS if all the numbers in the IS column are ≥ 0.

A coalition is PIS if the sum of all numbers in the IS column is ≥ 0

IS: Internal stability. PIS: Potential internal stability

Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

The overall worst performing coalition (WT) – both with respect to global payoff and global abatement – is also potentially internally stable. This is confirmed by closely examining the IS in table 5.7 above. No coalition, potentially internally stable or not, has lower global payoff or a lower level of global abatement than this coalition. It has an efficiency index of -0.1252 and an environmental index of -0.0534.

As we see from the table someone has to make side payments to Russia, India, Iran and Saudi Arabia to entice them to be part of the coalition. Japan, the Republic of Korea and Australia are willing to pay because they have negative benefit parameters, meaning that they actually benefit from increased emissions.

5.4 Non-internally stable coalitions

The best performing coalitions, including the social optimum, are not potentially internally stable. In fact, 1.5% of all possible coalitions outperform the best potentially internally stable coalition with regards to global payoff (BeffPIS), while approximately 3% of all possible coalitions result in a higher level of global abatement than the potentially internally stable coalition with the highest level of abatement (BenvPIS).

There are also as many as 404 coalitions (none of which are potentially internally stable) that result in a higher level of global abatement than the level of global abatement we get in the socially optimal outcome (where global payoff is maximized).

The coalition resulting in the overall highest level of global abatement (Benv) is presented in table 5.8 below. No coalition, potentially stable or not, has higher global abatement than this coalition. The coalition's environmental index is 1.1586 and the efficiency index is 0.9527.

The coalition has 14 members; only Japan and the Republic of Korea are not part of the coalition. As we can see the environmental index is larger than one; the coalition results in a higher level of global abatement than the socially optimal outcome. The coalition is not potentially internally stable.

		Payoff	Abatement	IS
	GLOBAL	4.597633	1.076167	-
1	**China**	**2.6776952**	**0.332261**	**0.434021**
2	**United States**	**1.8048946**	**0.098084**	**0.050215**
3	**European Union**	**3.1341173**	**0.050543**	**0.761101**
4	**Russian Federation**	**-0.9299623**	**0.286490**	**-1.421256**
5	**India**	**-0.4610915**	**0.131405**	**-0.708127**
6	Japan	-1.1110817	-0.000777	-
7	**Canada**	**0.0595108**	**0.007158**	**-0.039308**
8	Korea, Rep.	-0.4184661	-0.000109	-
9	**Iran, Islamic Rep.**	**-0.1712460**	**0.032463**	**-0.184555**
10	**Mexico**	**-0.0152440**	**0.011873**	**-0.066858**
11	**South Africa**	**-0.0250824**	**0.017327**	**-0.097116**
12	**Australia**	**-0.0735651**	**0.004839**	**-0.027608**
13	**Saudi Arabia**	**-0.1432268**	**0.025967**	**-0.147847**
14	**Brazil**	**0.1358811**	**0.006111**	**-0.031497**
15	**Ukraine**	**-0.1113816**	**0.060116**	**-0.325644**
16	**Indonesia**	**0.2458815**	**0.012416**	**-0.059055**

Table 5.8 – The coalition with the highest environmental index (Benv)

Countries marked in **bold** are members of the coalition

A coalition is IS if all the numbers in the IS column are ≥ 0

A coalition is PIS if the sum of all numbers in the IS column is ≥ 0

IS: Internal stability. PIS: Potential internal stability

Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

		Payoff	Abatement	IS
	GLOBAL	4.570955	1.075331	-
1	**China**	**2.659917**	**0.333507**	**0.417426**
2	**United States**	**1.798858**	**0.098452**	**0.044775**
3	**European Union**	**3.129295**	**0.050733**	**0.755246**
4	**Russian Federation**	**-0.942775**	**0.287564**	**-1.433003**
5	**India**	**-0.466941**	**0.131897**	**-0.713652**
6	Japan	-1.110219	-0.000777	-
7	**Canada**	**0.059126**	**0.007185**	**-0.039616**
8	Korea, Rep.	-0.418141	-0.000109	-
9	**Iran, Islamic Rep.**	**-0.172647**	**0.032585**	**-0.185943**
10	**Mexico**	**-0.015793**	**0.011918**	**-0.067365**
11	**South Africa**	**-0.025882**	**0.017392**	**-0.097855**
12	Australia	-0.045957	-0.000018	-
13	**Saudi Arabia**	**-0.144342**	**0.026064**	**-0.148958**
14	**Brazil**	**0.135487**	**0.006134**	**-0.031765**
15	**Ukraine**	**-0.114135**	**0.060341**	**-0.328189**
16	**Indonesia**	**0.245104**	**0.012462**	**-0.059609**

Table 5.9 – The coalition without Japan, the Republic of Korea and Australia (exJKA)

Countries marked in **bold** are members of the coalition

A coalition is IS if all the numbers in the IS column are ≥ 0

A coalition is PIS if the sum of all numbers in the IS column is ≥ 0

IS: Internal stability. PIS: Potential internal stability

Source: Holtsmark's spreadsheets (2013), based on my own calculations reported in table (4.4)

The coalition consisting only of members with positive benefit parameters (i.e. excluding Japan, the Republic of Korea and Australia) (exJAK) is presented in table 5.9 above. The efficiency index is 0.9446 and the environmental index is 1.1576. This coalition is interesting because the three excluded countries are the three countries with negative benefit parameters; this means that they actually benefit from emissions of CO_2 (see table 4.4). This should imply that a coalition without these three countries is able to abate more than the socially optimal coalition, where the three countries' benefits have to be considered when the coalition members maximize their joint payoff. As we can see from the environmental index, this coalition does lead to a higher global abatement than in the socially optimal coalition. It does not, however, result in higher global abatement than the coalition excluding only Japan and the Republic of Korea presented in table 5.8 (see discussion in the following section).

5.5 Interpretation and discussion

Table 5.10 provides a summary of the ten coalitions presented earlier in this chapter. Looking at the environmental index we see that three of the coalitions result in lower global abatement than the non-cooperative outcome. Two of the coalitions also result in higher abatement globally than the socially optimal outcome. While all of the 'ill-performing' coalitions are potentially internally stable (two of them are even internally stable without side payments), none of the coalitions abating more than the social optimum are.

Coalition	**Table**	**Global payoff**	**Global abatement**	I_{eff}	I_{env}	**IS**	**PIS**
WT[a)]	5.7	1.0426	0.1358	-0.1252	-0.0534	No	Yes
WIS[b)]	5.4	1.1859	0.1496	-0.0817	-0.0356	Yes	Yes
LIS[c)]	5.3	1.2857	0.1596	-0.0515	-0.0227	Yes	Yes
BaU[d)]	5.1	**1.4554**	**0.1772**	**0**	**0**	**Yes**	**Yes**
BIS[e)]	5.2	1.9586	0.2778	0.1526	0.1297	Yes	Yes
BeffPIS[f)]	5.5	4.3367	0.7266	0.8735	0.7081	No	Yes
BenvPIS[g)]	5.6	4.3151	0.7988	0.8670	0.8012	No	Yes
SO[h)]	5.1	**4.7537**	**0.9531**	**1**	**1**	**No**	**No**
exJKA[i)]	5.9	4.5710	1.0753	0.9446	1.1576	No	No
Benv[j)]	5.8	4.5976	1.0762	0.9527	1.1586	No	No

Table 5.10: Results for ten key coalitions

[a)] The worst performing coalition of them all (global payoff and global abatement), [b)] The worst performing internally stable coalition (global payoff and global abatement), [c)] An internally stable coalition with 11 members, [d)] The non-cooperative outcome, [e)] The best internally stable coalition (global payoff and global abatement), [f)] The best potentially internally stable coalition (global payoff only), [g)] The best potentially internally stable coalition (global abatement only), [h)] The socially optimal outcome, [i)] The coalition excluding only Japan, Korea, Rep. and Australia, [j)] The best coalition of them all (global abatement only)

I_{eff} : Efficiency index. I_{env} : Environmental index. IS: Internal stability. PIS: Potential internal stability

Sorted by the environmental index

Source: Holtsmark’s spreadsheets (2013), based on my own calculations reported in table (4.4)

Table 5.10 confirms that the best performing internally stable coalition (BIS) performs rather poorly in terms of relative global abatement, while the best performing potentially internally stable coalitions (BeffPIS and BenvPIS) both achieve a lot in terms of relative global abatement.

Holtsmark (2013) has done a similar analysis for the 16 world regions from the FUND model presented by Osmani and Tol (2009). While his results are similar to the results presented in this thesis, there were fewer both internally and potentially internally stable coalitions – with somewhat lower efficiency indexes (Holtsmark, 2013, p. 339).

Overall, the results of the analysis in this thesis are encouraging. Many of the possible coalitions are potentially internally stable. In addition, a lot of these coalitions perform rather well, both in terms of global payoff and in terms of global abatement (see figures 5.1 and 5.2). By allowing for side payments we get very close to the socially optimal outcome, both with regards to payoff and abatement.

The results for internally stable coalitions are less encouraging. While we do get larger internally stable coalitions when countries are heterogeneous than when they are homogeneous (recall that the largest internally stable coalition had only three members in the case with 16 identical countries), these large coalitions have lower gains – both in terms of global payoff and global abatement – than the best performing internally stable coalitions with only two or three members. A lot of the internally stable coalitions actually perform worse than the case with no cooperation; both the efficiency indexes and the environmental indexes are below zero.

The main trend for internally stable coalitions is that larger coalitions abate less than the small ones. This is consistent with the theoretical results presented in the paper by Pavlova and de Zeeuw (2012).

When the benefit of each individual country depends only on global abatement and the existing coalition is large, the contribution of one individual country matters little in the aggregate and each country has an increased incentive to free-ride on the abatement effort of others (i.e. to be non-members to an agreement). The coalitions with high gains of cooperation in terms of abatement are therefore not internally stable. This result is consistent with the claim Barrett (2003) makes; there is a trade-off between the depth (how well a coalition performs) and the breadth (how many members the coalition has) of cooperation.

However, internal stability is only one of the conditions a coalition must satisfy in order to be self-enforcing. I have in this thesis not considered whether or not any of the coalitions are externally stable. This condition imposes further restrictions on the set of attainable coalitions (i.e. coalitions that are self-enforcing). But, looking back at section 3.2, we see that the external stability condition is less restrictive than the internal stability condition for homogeneous countries. In fact, all coalitions with two or more members are externally stable when countries are identical.

While this result is not directly applicable to a situation where countries are different, at least it gives us an idea of what to expect. External stability implies that non-members have no interest in joining an existing coalition. That large coalitions are externally stable is consistent with the fact that countries are non-members because they want to free-ride on the abatement efforts of others. In already large coalitions the contribution of one additional country matters little in terms of global abatement, making it more likely that the country will benefit from simply staying out of the coalition, getting the benefit of the abatement effort of the coalition but not having to pay the cost of increased abatement itself.

Only Japan, the Republic of Korea and Australia should want to join a large coalition with high abatement; simply to reduce that coalition’s abatement level. They are in some cases willing to *pay* other countries to impose subsidies to emission generation. Their benefit parameters are negative, so they actually gain from decreased abatement. As we saw from the results presented above, the coalition excluding Japan, the Republic of Korea and Australia actually abates more than the socially optimal coalition. This is because when a coalition maximizes their aggregated payoff with respect to abatement, the inclusion of a country with a negative benefit parameter reduces the optimal abatement level for the coalition as a whole.

It is, however, interesting to note that the coalition that results in the highest level of abatement overall (see table 5.8) actually includes Australia (though not Japan and the Republic of Korea). While Australia itself does not profit from higher abatement, Australia’s benefit parameter is close to zero (see table 4.4) and when their joint payoffs are maximized the gain the other coalition members recieve from high abatement outweighs Australia’s small individual loss. This coalition is not potentially internally stable; even though the coalition's joint payoff is maximized, the gains of cooperation are not large enough to compensate the countries that do not gain from being part of the coalition.

It is also interesting to note that some coalitions (whether they are (potentially) internally stable or not) abate less and result in a lower global payoff than the non-cooperative outcome because there are countries with negative benefit parameters in the model. All coalitions with negative efficiency and environmental indexes include at least one of those three countries. The worst internally stable coalition, both in terms of global abatement and global payoff (WIS), includes Japan and the Republic of Korea (see table 5.4), while the worst performing coalition of them all, also in terms of both global abatement and global payoff (WT), includes all three countries as members (see table 5.7). This coalition is potentially internally stable; Japan, the Republic of Korea and Australia make side payments to the other four members to entice them to be part of this 'bad' coalition (see table 5.6).

While Japan, the Republic of Korea and Australia are the main contributors to the bad results, China, the US and the EU contribute to most of the well-performing coalitions. The best performing internally stable coalition in terms of both global payoff and global abatement (BIS) consist of only China and the US (see table 5.2). This implies that negotiations between only these two countries could yield good results.

China and the EU are members of the best performing potentially internally stable coalition in terms of payoff (BeffPIS). They make side payments to other members to entice them to be part of the coalition (see table 5.4). The potentially stable coalition resulting in the highest level of abatement (BenvPIS) includes also includes China, the US and the EU. These three make side payments to other members of the coalition (see table 5.5).

Interestingly, Japan also makes side payments to the other members in both of these coalitions (and the Republic of Korea does so in the BenvPIS coalition). This implies that while these coalitions are the best performing (potentially internally stable) coalitions, Japan (and the Republic of Korea) is actually paying the other members in the coalition to abate less than they would have if Japan was not a member of the coalition.

While side payments allow for much larger coalitions with relatively high both global abatement and global payoff by making recipients more inclined to become members, they make the donors worse off. In addition to making the side payments less credible, there may also be some moral complications if the countries receiving the side payments are financially better off than the donors. However, according to the results presented in this thesis, the countries making the side payments are actually what we consider to be relatively rich countries. China, the US, the EU and Japan are countries making side payments, while countries such as India, South Africa and Indonesia are the recipients.

6 Conclusion

The purpose of this thesis has been to calculate the cost and benefit parameters in a linear benefit/quadratic cost model of international climate negotiations for 16 individual countries and solve the model in order to analyze the stability properties of the possible coalitions.

The European Union and the 15 specific countries were chosen because they had the highest emissions of CO_2 in 2005. The parameters were calculated from 16 world region parameters presented in a paper by Osmani and Tol (2009), and the model was solved using a spreadsheet developed by Bjart Holtsmark (2013). This is a partial analysis, where the terms global and social optimum apply to a world consisting only of these 16 countries.

Holtsmark's spreadsheets (2013) calculate the payoffs and abatement level for each country in each of the 65 519 possible coalitions assuming that the coalition countries agree to maximize their joint payoff, as well as the outcome where no countries cooperate. The spreadsheets also calculate the global payoff and the global abatement for each coalition and rate the coalitions according to global payoff and global abatement – relative to global payoff and global abatement in the two extremes where no one cooperates and where everyone cooperates – using two indexes; the efficiency index (which measures global payoff) and the environmental index (which measures global abatement). In addition, the spreadsheets reveal which coalitions are internally and potentially internally stable.

One result of this calculation is that relatively few coalitions are internally stable (without side payments) but that the majority of the coalitions are potentially internally stable when side payments are introduced.

The internally stable coalitions generally achieve little, both in terms of relative global payoff and relative global abatement. In addition, the smaller internally stable coalitions clearly outperform the larger ones – both with respect to relative global payoff and relative global abatement.

How does this translate to the real world? The results on internal stability imply that small coalitions perform better than large coalitions (in terms of both global abatement and global payoff). The best performing internally stable coalition is a coalition consisting only of China and the US. Hence, perhaps it would be optimal to let China and the US negotiate alone, rather than to aim for a broad coalition worldwide.

However, if the possibility for side payments is included, the picture is much brighter. Many of the potentially internally stable coalitions actually do quite well. While none outperform the grand coalition (were all countries cooperate and maximize their joint payoff) some are quite close. The reason is that when side payments are allowed, the countries that gain from cooperation are able to compensate countries that would otherwise loose and not be part of the coalition. This increases the set of attainable coalitions, and by cooperating the countries are able to achieve a lot in terms of relative global abatement.

Still, some questions remain. While internal stability is arguably the most ‘important’ stability condition – keeping countries outside a coalition is easier than forcing countries to join – it would be interesting to find out which of the 65 519 coalitions are externally stable. It would also be interesting to see how the stability of the coalitions change if the benefit parameters of Japan, South Korea and Australia, no longer take negative values. Would we perhaps get fewer but better performing internally stable coalitions? The grand coalition itself should at least result in a higher level of abatement.

It would also be interesting to be able to say something about the abatement levels in more absolute terms. While I claim that many potentially internally stable coalitions perform well relative to the grand coalition, I do not make any absolute claims about the level of abatement in the grand coalition itself. While the grand coalition is the coalition that maximizes global payoff with respect to abatement,[29] the resulting optimal level of abatement may or may not be able to reduce the risk of global warming.

---

[29] The objective of a coalition in this thesis is to maximize their joint payoff with respect to abatement. Global abatement is the sum off all countries’ payoffs (whether they are part of the coalition or not).

Bibliography

TABLES, THE CLIMATE FRAMEWORK FOR UNCERTAINTY, NEGOTIATION AND DISTRIBUTION (FUND), VERSION 2.8. (2006). Retrieved June 25, 2012, from http://www.fund-model.org/versions.

Barrett, S. (2003). *Environment & Statecraft.* Oxford University Press.

Barrett, S. (2006). Climate Treaties and "Breakthrough" Technologies. *The American Economic Review, Vol. 96, No.2*, pp. 22-25.

d'Aspremont, C., Jaquemin, A., Gabszewicz, J., & Weymark, J. (1983). On the stability of collusive leadership. *Canadian Journal of Economics 16*, pp. 17-25.

Finus, M. (2004). Stability and design of international environmental agreements: the case of transboundary pollution. In H. Folmer, & T. Tietenberg, *The International Yearbook of Environmental and Resource Economics 2003/2004* (pp. 82-158). Edgar Elgar Publishing.

Gibbons, R. (1992). *A Primer in Game Theory.* Pearson Education.

Handmark, Inc. (2011, October 5). Economics - Oxford Dictionary for iPhone. (Version 2.0.1). Oxford University Press.

Hansen, R. (2007). *Pristeori.* Oslo: BI Forlag.

Hoel, M. (2005). Global Warming and Other Transboundary Environmental Problems. In O. Bjerkholt, & X. Ping, *Frontier Environmental Issues* (pp. 37-64). Fudanpress.

Holtsmark, B. (2013). International cooperation on climate change: Why is there so little progress? In E. Elgar, *Handbook on Enery and Climate Change (in press)* (pp. 327-343).

IPCC. (2007). *The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the International Governmental Panel on Climate Change.* Cambrigde, UK and New York, USA: Cambridge University Press.

Osmani, D., & Tol, R. (2009). Toward Farsightedly Stable International Environmental Agreements. *Journal of Public Economic Theory, 11 (3)*, pp. 455-492.

Pavlova, Y., & Zeeuw, A. d. (2012, July 25). *Asymmetries in international environmental agreements.* Retrieved November 22, 2012, from Environment and Development Economics: Available on CJO2012 doi:10.1017/S1355770X12000289

Perman, R., Ma, Y., Common, M., Maddison, D., & McGilvray, J. (2011). *Natural Resource and Environmental Economics* (4th ed.). Addison Wesley.

Sandler, T. (1997). *Global Challenges - an approach to environmental, political and economic problems.* Cambridge University Press.

Stern, N. (2006). Executive Summary. In N. Stern, *The Stern Review: The Economics of Climate Change* (pp. i-xxvii). London: HM Treasury.

Stiglitz, J. E., & Walsh, C. E. (2006). *Principles of Microeconomics* (Fourth ed.). New York - London: W. W. Norton & Company.

Sydsæter, K. (2000). *Matematisk analyse - bind 1.* Gyldendal Norsk Forlag AS.

Tol, R. (2006, April 10). *THE CLIMATE FRAMEWORK FOR UNCERTAINTY, NEGOTIATION AND DISTRIBUTION (FUND), TECHNICAL DESCRIPTION, VERSION 2.8.* Retrieved June 25, 2012, from http://www.fund-model.org/versions.

Watson, J. (2002). *Strategy - An Introduction to Game Theory.* W. W. Norton & Company.

World Bank, the. (2005a). *CO2 emissions (kt)*. Retrieved June 22, 2012, from http://data.worldbank.org/indicator/EN.ATM.CO2E.KT?page=1

World Bank, the. (2005b). *Population, total*. Retrieved June 22, 2012, from http://data.worldbank.org/indicator/SP.POP.TOTL?page=1

World Bank, the. (2005c). *GDP (current US$)*. Retrieved June 22, 2012, from http://data.worldbank.org/indicator/NY.GDP.MKTP.CD?page=1