The Stability and Effectiveness of International Climate Agreements: The Role of Carbon Trade, Bargaining Power and Enforcement

Shumin Yu

Thesis committee

Promotor
Prof. Dr E.C. van Ierland
Professor of Environmental Economics and Natural Resources
Wageningen University & Research

Co-promotors
Dr H.-P. Weikard
Associate professor, Environmental Economics and Natural Resources Group
Wageningen University & Research

Dr X. Zhu
Associate professor, Environmental Economics and Natural Resources Group
Wageningen University & Research

Other members
Prof. Dr A.G.J.M. Oude Lansink, Wageningen University & Research
Prof. Dr J. Gupta, University of Amsterdam
Dr C. Hagem, Statistics Norway, Oslo, Norway
Dr H. Houba, VU Amsterdam

This research was conducted under the auspices of the Wageningen School of Social Sciences (WASS).

The Stability and Effectiveness of International Climate Agreements: The Role of Carbon Trade, Bargaining Power and Enforcement

Shumin Yu

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Thursday 20 April 2017
at 11 a.m. in the Aula.

Shumin Yu
The Stability and Effectiveness of International Climate Agreements: The Role of Carbon Trade, Bargaining Power and Enforcement,

147 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands [2017]
With references, with summary in English

ISBN 978-94-6343-064-7
DOI: 10.18174/402826

Contents

Chapter 1

Introduction

1.1Background

Climate change is a serious threat to the welfare of the current and future generations. Although efforts from international communities have been made to decelerate the process of climate change, global average atmospheric temperature has increased and will continue to increase for decades even under stringent climate policies. According to the latest Fifth Assessment Report (AR5) by IPCC, many of the observed changes in the climate system are unprecedented. The globally averaged surface temperature has shown an increase from 0.65℃ to 1.06℃ over the period 1880 to 2012 (IPCC 2014). As a direct result of global warming, glaciers and ice sheets have been disappearing at increasing rates over the period 1992 to 2011 (IPCC 2014). As a consequence, the global mean sea level has already risen by 0.19 m over the period 1901 to 2010 (IPCC 2014). Moreover, if fossil fuel stocks will get depleted, the ultimate impacts on sea levels could be catastrophic for many countries (Winkelmann et al. 2015). There is now also overwhelming evidence that global warming is the result of anthropogenic greenhouse gas (GHG) emissions (IPCC 2014). Emissions of CO_2 from fossil fuel combustion and industrial processes contributed about 78% to the total GHG emissions increase from 1970 to 2010 (IPCC 2014). Hence, mitigating anthropogenic CO_2 emissions is crucial to address the problem of global warming. Particularly, due to the irreversibility of climate change (Carraro et al. 2006), strategies of CO_2 emissions mitigation are needed to reduce the negative impacts of climate change.

GHG emissions generate negative externalities that are not confined to the borders of an emitting country. It should be described as a transboundary pollution problem and should be addressed through cooperation at an international level. The optimal GHG abatement policy internalises negative externalities on other countries. The other side of the same coin is that benefits from emissions mitigation of one country are positive externalities for other countries as they can be shared by others. Due to the public good properties of GHG mitigation (i.e. non-excludability and non-rivalry), each country has incentives to free-ride on others' mitigation. These free-riding incentives are an obstacle for the cooperative action on GHG emissions mitigation (Barrett 1994).

Barrett (1990) points out that one way out of the public goods dilemma on emissions mitigation is to form agreements between sovereign countries. Such agreements involve mutual consent to a global climate policy and its implementation. However, countries are characterised by heterogeneous mitigation costs and benefits. These asymmetries imply that benefits of cooperation differ between countries and there is no straightforward way to allocate the costs in an equitable way. Moreover, cooperating countries who are "doing the job" could be worse off compared to a non-cooperating country. Due to the lack of a supranational authority to enforce participating countries to fulfil their mitigation commitments, agreements have to be implemented voluntarily by countries. To facilitate the formation of international climate agreements, effective economic and institutional instruments need to be designed and implemented. The core of this thesis is the study of design features of an international climate agreement that improve its effectiveness and acceptance by sovereign countries.

1.2 International climate agreements

International climate agreements (ICAs) are formed between sovereign countries and designed with the aim of reducing GHG emissions. The first international climate agreement is the United Nations Framework Convention on Climate Change (UNFCCC) which was signed in 1992. Until now 194 countries have been involved in this treaty. UNFCCC (1992) sets its objective as to "stabilise greenhouse gas concentrations in the atmosphere at a level that will prevent dangerous human interference with the climate system". However, mitigation targets for member countries in UNFCCC are non-binding, which means that UNFCCC cannot enforce the reduction of GHG emissions. Nevertheless, as the starting point for ICAs, the UNFCCC shows the potential to tackle the problem of climate change in a cooperative way. Following the objective of the UNFCCC to reduce GHG concentrations, the Kyoto Protocol was signed in 1997. As one step further in the UNFCCC process, the Kyoto Protocol sets legally binding mitigation targets for selected members in a stepwise method. Furthermore, an enforcement mechanism has been included to make signatories fulfil their commitments. To increase the cost-effectiveness of emission reductions, market-based mechanisms are designed in the protocol. However, the Kyoto Protocol failed to include binding targets for developing countries. Clearly, the Kyoto Protocol has a limited ability to induce significant reductions of worldwide GHG emissions. Facing threats from climate change, a stable post-Kyoto ICA with high mitigation efficiency in reducing global GHG emissions is urgently needed. The Paris Agreement, which was adopted in December 2015 and went into effect on 4 November 2016, is considered as a historic agreement for climate change mitigation. To limit the global average temperature increase below 2 ℃, the Paris agreement seeks participation of all parties to the UNFCCC. So far, the two largest GHG emitters, China and the United States, have ratified this agreement.

The public good nature of GHG emissions mitigation implies that individual mitigation incurs national costs but generates commonly shared benefits. Thus, each individual country has potential incentives to free ride on other's mitigation efforts. Given this, there are two general aspects that should be considered in the design of an effective ICA.

Firstly, due to the lack of a supranational authority to enforce cooperative behaviours, ICAs have to be self-enforcing. This implies that countries must have sufficient incentives to voluntarily participate in an ICA. The economic literature on ICA formation based on non-cooperative game theory (Barrett 1992, 1994; Hoel 1992; Carraro and Siniscalco 1993) shows that self-enforcing ICAs are characterised by a small size. ICAs with a larger size can only be reached when the gains from cooperation are small. To improve countries' incentives to participate, some studies focus on reducing free-riding incentives by using various transfer schemes (e.g. Barrett 2001; Eyckmans and Finus 2006; Weikard et al. 2006), and applying different membership rules (Finus et al. 2005; Finus 2008) or social norms (Lange and Vogt 2003; van der Pol et al. 2012). To reduce free-riding incentives that result in countries' noncompliance with their commitments, the role of enforcement mechanism comprising monitoring and punishment in enhancing countries' compliance levels is discussed and studied in some literature (Hovi and Areklett 2004; Barrett 2008; McEvoy 2013).

The second aspect we need to consider for ICAs design is the efficiency of ICAs, which includes the cost effectiveness and the mitigation effectiveness. From an economic perspective, cost effectiveness is achieved by implementing the mitigation targets of an ICA at minimum costs. It is commonly acknowledged that mitigation costs can be minimised through market-based instruments, for example through carbon trade (Montgomery 1972; Tietenberg 1985; Barrett 1992; Helm 2003). Mitigation effectiveness concerns the "depth" of ICAs. The most frequently employed approach to addressing mitigation effectiveness of ICAs has been setting the mitigation standards and regulating countries' emissions levels, the so-called command-and-control instrument, whereby a high mitigation level can be achieved by setting mitigation targets (Sinclair 1997; Endres and Finus 2002).

1.2.1 Carbon trade

The approach of marketable emission permits or carbon trade has proven its cost-effectiveness advantage in reducing GHG emissions (Montgomery 1972; Tietenberg 1985; Barrett 1992; Rose and Stevens 1993; Leiby and Rubin 2001). Due to the heterogeneity in marginal abatement costs and benefits between countries, trade in emission permits enables carbon buying countries to reduce abatement costs and generates revenues for carbon selling countries. In addition to the cost-efficiency improvement, carbon trade can also alleviate the impact of free-riding incentives by serving as a transfer scheme to redistribute coalitional gains.

While carbon trade implements abatement targets at minimal costs, it has proven to be difficult to establish consensus on binding targets. Moreover, for the design of a carbon trade mechanism, it is also essential to elaborate on how to allocate initial emission permits among countries that participate in carbon trade. The initial permits allocation determines the cost-savings or revenues of carbon traders and thus their incentives to trade. Free allocation of initial emission permits was a standard feature of most existing carbon trade schemes, for example the European Union Emissions Trading System (EU ETS). Generally, there are two allocation schemes. One scheme allocates the initial emission permits according to some given criteria. For example, taking into account the acceptability, some pragmatic rules based on sovereignty features like GDP, population, or historical emissions are proposed (Larsen and Shah 1994; Bohm and Larsen 1994; Edmonds et al. 1995). Considering the equity concept of burden-sharing in GHG emissions reduction, some equity rules based on the normative criteria are introduced as an improvement to pragmatic rules (Rose and Stevens 1993; see Weikard 2004 for a more fundamental discussion). Both pragmatic and equity rules can be characterised as exogenous permit allocation rules. Alternatively, initial permits can be allocated endogenously. Helm (2003) suggested an endogenous allocation rule based on a strategic choice mechanism. According to this rule, initial permits and the after-trade emission levels are chosen non-cooperatively based on individual welfare maximisation. The endogenous permit allocation rule improves the cost efficiency in emission permits allocation across carbon traders.

As to the effect of the emission permit allocation rule on countries' participation incentives in ICAs, Altamirano-Cabrera and Finus (2006) examined the impact of permit trading on the stability of international climate agreements in a game theoretical model by applying exogenous permit allocation rules. Their simulation results show that permit trading based on pragmatic allocation rules is more successful in improving the stability and abatement efficiency of ICAs, as compared to the one based on equitable allocation rules. To study the effect of the endogenous permits allocation scheme on participation incentives, Carbone et al. (2009) formulated a two-stage coalition formation model that analyses the formation and abatement efficiency of an international emission trade agreement based on Helm's endogenous permits choice model. They have shown that an international climate agreement with international carbon trade (and, hence, cost efficient abatement) can be formed between developed and developing countries under endogenous permits choice. In their model countries with lower marginal abatement costs are incentivised to join carbon trade by seeking trade revenues. Their participation improves the coalitional abatement level and the cost-efficiency of abatement. In both analyses of Carbone et al. (2009) and Helm (2003), emission permits allocated to all carbon traders are not constrained. Without constraints on the total or individual permit choices, as studied by Helm (2003), the motivation to raise revenues from carbon trading could result in inefficiencies due to excessive permit choices, so-called 'hot-air'. Another problem that could arise in such an unconstrained carbon market concerns the stability. Standard climate agreements are characterised by open membership, which implies free entry and exit. For an unconstrained carbon market with open membership, large revenues generated from carbon trading could induce excessive participation of countries with lower marginal abatement costs. The result could be that emission permits are oversupplied, which could undermine the stability of a carbon market through participation incentives that originate from selling carbon permits. Therefore, imposing a constraint on permit choices might not only mitigate the hot-air effect, but may also help to stabilise a carbon market by avoiding excessive participation. Hence, it is worth studying how exogenous constraints on allowance choices change the incentives to participate in an international carbon market and its environmental effectiveness.

In the Kyoto Protocol, the Clean Development Mechanism (CDM) is designed to reduce members' free-riding incentives and improve the abatement effectiveness of the protocol by lowering signatories' abatement costs. Under the CDM, two different kinds of carbon traders are involved. One is the signatory of the Kyoto protocol (Annex B countries), the others are non-signatories. Signatories are allowed to earn tradeable emission reduction credits by implementing mitigation projects in the outsider developing countries. This carbon trade option raises questions about its impact on the incentives to join and the effectiveness of an

ICA for mitigation. There are two coalitions that can be formed: a climate coalition for GHG emissions mitigation and a coalition for carbon trade. When these two coalitions are considered in parallel, the mitigation coalition and the non-signatories are considered as potential carbon traders and allowed to join the carbon market. As compared to a simple mitigation coalition, incentives with an open carbon market will be different in the sense that countries are not only incentivised to join an ICA for obtaining cooperation gains, but are also motivated to join the carbon market for trade benefits. Furthermore, an open carbon market not only increases the welfare of signatories of the ICA, but also of non-signatories. If the latter effect from trade benefits is sufficiently strong, then leaving the coalition will become more attractive and the carbon market has a negative effect on participation incentives and the success of the ICA. This can be translated into the research problem of how players' incentives to participate in an ICA change if an independent carbon market is established outside of an ICA. It might offer an alternative or complementary policy instrument to facilitate mitigation. Furthermore, facing the advantage of low abatement costs and trade revenues, the question is whether or not the mitigation coalition is willing to join the carbon market. This concerns the problem of the interplay of a mitigation agreement and a carbon market. In this thesis, the impact of a carbon market as an independent party on the stability and effectiveness of an ICA for GHG emissions mitigation is investigated.

1.2.2 Nash bargaining solution

The Samuelson rule defines how cooperation is most efficient for public good provision. With respect to the climate mitigation problem, international cooperation on GHG emissions mitigation generates positive net gains compared to the non-cooperative Nash behaviour. The potential positive gains from cooperation show the need for negotiation between countries with free-riding incentives. For example, some international climate agreements (e.g. the UNFCCC, the Kyoto Protocol) are established through multilateral negotiations. Negotiation between players for collaborative benefits can be regarded and solved as a bargaining problem (Nash 1950). Hence, bargaining is a tool to distribute collaborative gains and to facilitate cooperation by providing well-designed sharing schemes for collective gains. Nash (1950) provided an axiomatic approach to distributing collective gains in the context of cooperative bargaining theory. Under this approach there is a unique solution to a bargaining problem, which is known as the Nash bargaining solution satisfying a set of axioms, such as individual rationality and Pareto efficiency (see Nash 1950). The Nash bargaining solution can obtained by solving a maximisation problem and has been widely applied to distributional problems (see Grout 1984; Jackson and Moulin 1992; Han et al. 2005; Carraro et al. 2006).

However, so far only few studies have applied the Nash bargaining solution to the problem of international climate cooperation. To examine the impacts of different sharing schemes in coalitional payoffs sharing, Carraro et al. (2006) applied the Nash bargaining solution to coalitional surplus sharing in a climate coalition. In their model, coalition members are assumed to be symmetric with equal bargaining power. However, in real climate negotiations negotiating countries are asymmetric. This asymmetry impacts countries' bargaining power, which reflects each negotiator's ability to obtain the share of cooperative gains. Thus, the bargaining outcome on the distribution of collective gains is subject to bargaining power. Then the key issue of bargaining over gains from international climate cooperation is to identify what constitutes the bargaining power of each negotiator. Therefore, it is interesting to explore potential sources that could induce differences of negotiators' bargaining power and identify the determinants of bargaining power in international climate negotiations. Furthermore, a bargaining solution is not just affected by bargaining power, but also by players' outside options (Wagner, 1988; Muthoo, 1999; Powell, 2002). The outside option payoffs impose a lower bound on the bargaining solution since no one needs to accept an agreement where he is worse off compared to what he can obtain otherwise. Therefore, it is also interesting to study the impact of the Nash bargaining solution applied to gains from climate cooperation and considering outside options on the incentives to join an international climate agreement in the first place.

1.2.3 Enforcement

Free-riding incentives exist in both the participation and implementation stage of ICAs. The free-riding incentives in the former case can be overcome by some institutional designs considering membership rules (Finus et al. 2005) or side payments (e.g. Eyckmans and Tulkens 2003; Carraro et al. 2006; Eyckmans and Finus 2006; Weikard et al. 2006). To improve the success of an ICA, an enforcement mechanism is needed to incentivise signatories to comply with their commitments. The idea that contracts need enforcement is not new (e.g. Buchanan 1975). For the case of ICAs, Barrett (2008) analyses the features of the Kyoto protocol and points out that enforcement mechanisms are essential and imperative to ensure the effectiveness of ICAs. Moreover, Finus (2008) reviews the enforcement mechanism adopted by the Kyoto protocol and confirms positive effects of the enforcement, but also proposes measures to improve the enforcement mechanisms of the Kyoto Protocol. As a further step, Hovi et al. (2012) formulate a pragmatic and credible compliance enforcement system for post-Kyoto climate agreements. With respect to the impact of the enforcement on the participation in and compliance with self-enforcing ICAs, McEvoy and Stranlund (2009) study a coalition formation game that includes an enforcement system with costly monitoring. Their theoretical results show that under costly enforcement the set of stable ICAs is smaller, but stable coalitions can reach higher levels of participation and abatement compared to costless enforcement where compliance is taken for granted. Based on the same game structure, McEvoy and Stranlund (2010) study the effect of costly enforcement on the efficiency of voluntary environmental agreements. They find that a voluntary environmental agreement can be more efficient in reaching emissions targets than an emissions tax under the condition that the agreement is enforced by a third party that is financially supported by the members of the agreement. Their results also imply that free-riding incentives can be reduced if signatories bear enforcement costs. The analyses of McEvoy and Stranlund (2009, 2010) provide insights into the design of enforcement mechanisms in ICAs, where an effective enforcement should be undertaken by an independent third party and funded by all signatories.

However, McEvoy and Stranlund (2009, 2010) only consider the case of full compliance by all members, which results in abatement decisions are restricted to whether or not to control emissions. Therefore, their assumption does not consider partial enforcement that may result from signatories' choice of enforcement expenditures. Signatories of ICAs are motivated by cooperative gains to support the enforcement, however they will not contribute an amount that is beyond their compliance benefits. If their contribution to enforcement is not sufficient for full compliance, certain amounts of noncompliance may occur. The social optimum is reached by full compliance of signatories in an ICA. When enforcement is costly, however, the optimal enforcement could induce the partial compliance (Arguedas 2005, 2008; Stranlund 2007).

By allowing for partial compliance on the formation of ICAs, one chapter of this thesis investigates the optimal enforcement policy in an ICA. This enforcement policy involves the setting of the optimal abatement target and monitoring expenditure. Furthermore, the impact of an optimally designed enforcement mechanism on incentives to participation and compliance in a self-enforcing climate coalition is also studied in that chapter.

1.3 Objective and research questions

The aim of this thesis is to analyse the formation of climate coalitions and to gain insight into the implementation mechanisms of ICAs in an game theoretic framework. Introducing institutional design elements of ICAs, this thesis aims to identify optimal outcomes where everyone gains in an incentive compatible environment. In order to achieve this objective, the following research questions are addressed:

- What is the impact of a carbon market on regional incentives to join an ICA for GHG emissions mitigation when the carbon market is established independently of this agreement?
- How does an individual allowance choice constraint impact the stability and effectiveness of a carbon trade agreement?
- What is the impact of using the Nash bargaining solution for distributing coalitional gains under different sets of bargaining weights on the stability and effectiveness of international climate agreements?
- How can an optimal enforcement mechanism for an ICA be designed, and what is the impact of an optimally designed enforcement mechanism on participation and compliance?

1.4 Methodology

To address the research questions, game theoretical models are solved analytically and by numerical analysis. Negotiations on international climate cooperation are essentially strategic interactions between countries with different interests. Therefore, game theory is an appropriate theoretical instrument to analyse negotiating countries' behaviours and incentives. The problem of international climate cooperation is characterised by the lack of property rights, externalities and the absence of a supranational authority to enforce climate policies (Folmer and van Mouche 2000; Schmidt 2000). These characteristics explain why non-cooperative game theory is often used as a theoretical foundation to analyse the international cooperation on climate mitigation. By applying a game theoretical approach, the formation of ICAs is modelled as a coalition formation game, whereby the participation incentive is analysed. In such a context, a stable climate coalition is a Partial Agreement Nash equilibrium (PANE) (see Chander & Tulkens 1995). In a PANE, a coalition chooses actions based on joint payoffs maximisation while the outsiders to the coalition choose actions that maximize their individual payoffs. In this thesis, I use this game theoretic concept. By applying backward induction, the subgame perfect Nash equilibrium is obtained as the solution concept of such a game. Under subgame perfect Nash equilibrium, the individual rationality of signatories at each stage is ensured.

For the numerical analysis in this thesis, I employ a calibrated numerical model – an adjusted version of the STAbility of COalitions model (STACO version 3.0) as documented by Dellink et al. (2015). STACO is an integrated assessment model. It links a (simple) climate change module with a game theoretic module for economic and policy analysis. STACO formulates a two-stage cartel (ICA) formation game for twelve heterogeneous regions in the world. STACO model provides a basis for addressing the research questions in this thesis with respect to the stability and performances of ICAs in the context of heterogeneous players. In the two-stage game of STACO, regions decide at the first stage whether or not to join a coalition. There are two strategies to be chosen, either to join or not to join. At the second stage, abatement strategies are chosen simultaneously by coalition members and outsiders. For coalition members, their abatement is chosen cooperatively considering the abatement levels of outsiders as given, whereas the outsiders choose their abatement non-cooperatively considering all others' abatement as given. This two stage-game in STACO implies that the membership choice is once and for all and cannot be changed in the later periods of the time horizon defined in STACO.

In order to capture the long-term effect of climate change, STACO adopts a time horizon of 100 years. The climate module of STACO relates GHG emission paths to GHG concentrations and atmospheric temperature change for the whole time horizon. In STACO, the GHG emissions specifically focus on the CO_2 emissions. It is worth noting that numerical results in Chapters 2, 3 and 4 of this thesis are based on the specific modelling approach and the calibration in the STACO model. Therefore, the results for coalition stability and performances would be different from the ones derived from other integrated assessment models, for example the RICE model (Nordhaus and Yang 1996), the ClimNeg World Simulation (CWS) model (Bréchet et al. 2011; Eyckmans and Tulkens 2003; Eyckmans and Finus 2006), the Model of International Climate Agreements (MICA) (Lessmann et al. 2009; Lessmann and Edenhofer 2011) and the WITCH model (Bosetti et al. 2006). In particular, in the STACO model the functional forms of abatement benefits and costs are specified as cubic and linear forms, respectively. For a full specification of the latest version of the STACO model (STACO 3) the reader is referred to Dellink et al. (2015).

More complex game theoretical models are studied in this thesis, and the STACO model is adapted for the numerical analysis of these models. For the first research question, I extend the two-stage game to four stages by including the formation of a carbon market. To solve the second research question, I modify the original model to the formation of a carbon market with constrained endogenous allowance choices, where allowances, abatement levels and carbon prices are generated endogenously in an international carbon market system. To answer the third research question, I introduce the Nash bargaining solution with different sets of unequal bargaining weights into a climate coalition to redistribute the cooperative gains across coalition members. The fourth question is addressed with an analytical model.

1.5 Outline of the thesis

The remaining of the thesis consists of four main chapters which answer the research questions described in section 1.3.

Chapter 2 analyses the impact of an independent carbon market with endogenous allowance choice on the incentive structure of an international climate coalition for GHG emissions mitigation. I formulate a four-stage game theoretical model, in which the climate coalition and the carbon market are formed sequentially. To show the differences in participation incentives and performances from the two-stage cartel formation game, two scenarios are examined by a simulation analysis.

Chapter 3 investigates the formation and mitigation efficiency of an international carbon market where the individual choice of emission allowances is constrained by imposing an exogenous constraint. To show the impact of the constraint on the choices of allowance and abatement level, I analyse the strategy of carbon traders theoretically and numerically through increasing the strictness of constraints.

Chapter 4 explores the application of the Nash bargaining solution with different sets of asymmetric bargaining weights to the distribution of cooperative gains in a climate coalition. I identify the possible determinants of bargaining powers for the negotiators in a climate negotiation, and provide different sets of bargaining weights used for the distribution of coalitional gains.

Chapter 5 studies the design of an optimal enforcement mechanism in an ICA and the impact of this enforcement mechanism on the participation and compliance level of an ICA.

The Chapter 6 summarises the answers to the research questions, and contains the synthesis and conclusions.

Chapter 2

International climate agreements and the scope for carbon trade with endogenous permits*

In this paper, we develop an international climate policy game with a mitigation agreement and a carbon market that is open for all. The carbon market is modelled following Helm's (J Public Economics, 2003) suggestion that participants may freely choose their initial permit endowments and then trade. In a game theoretic model we explore incentives for market participation, incentives to join a mitigation agreement and the interlinkages between the two. We employ a numerical analysis of a 7-region version of the STAbility of COalitions (STACO) model. We find that number and size of stable mitigation coalitions are smaller with than without a carbon market. The mitigation coalition has no incentive to join a carbon market with non-signatories, since the benefits from free riding surpass the net gains from carbon trade. However, some non-signatories would join a carbon market which helps to reduce global emissions.

* This chapter is based on a submitted manuscript and a working paper presented in a thematic session of the 5th World Congress of Environmental and Resource Economists (WCERE) in Istanbul, June 28-July 02, 2014: Yu, S., Weikard, H.-P., Zhu, X., van Ierland, E.C. (2016). International climate agreements and the scope for carbon trade with endogenous permits.

2.1 Introduction

As it remains difficult to reach an effective global climate agreement, it has been suggested to rely on unilateral actions or move forward with a partial coalition that would facilitate the formation of an International Climate Agreement (ICA) (Carraro and Siniscalco, 1993; Brandt, 2004). A prominent tool to meet mitigation targets, whether unilateral or global, are carbon emission trading schemes. The European Union (EU), for example, has introduced the European Union Emissions Trading System (EU ETS), which is a major tool within the EU to reach the mitigation targets of the second commitment period of the Kyoto protocol. As the EU is developing climate policies for the time beyond 2020, the scope for extending the European carbon market towards a global market has received increasing attention (e.g. Behr et al., 2009). The Clean Development Mechanism (CDM) defined in the Kyoto protocol allows Annex B countries that have accepted a cap on emissions to earn certified emission reductions by implementing mitigation projects in developing countries. Because these certified emission reductions may be traded in an emission trading scheme, the CDM can also be interpreted as a step in the direction of a global carbon market.

In this paper we examine the stability of ICAs for mitigation when at the same time, but independently, a carbon market offers an alternative or complementary policy instrument to facilitate mitigation. Our analysis extends the frequently-used two-stage game for the analysis of international environmental agreements (e.g. Carraro and Siniscalco 1993). We amend the game with a carbon market that is open to all. Following a suggestion by Helm (2003) we model a carbon market with endogenous emissions permit choice. Upon entering the market each country makes an announcement of its emissions cap, i.e. its initial endowment with permits. Our game theoretic analysis sheds light on the interplay of a mitigation agreement and a carbon market. More specifically, we examine the sequential formation of two agreements, one on mitigation and one on carbon trade. Our game has four stages. At stage 1, players make their decisions on whether or not to join the mitigation agreement. We refer to those who join as the coalition members, and the remaining players are referred to as singletons. From stage 2 onward coalition members coordinate their climate policies and act as a single player. At stage 2, the coalition and the remaining singletons decide whether or not to join the carbon market. At stage 3, market participants announce their initial permit endowments. At stage 4, market participants can sell or buy permits and all players choose their mitigation levels. We solve the game backwards employing Sub-game Perfect Nash

Equilibrium. The equilibrium determines membership in the mitigation agreement, carbon market participation, carbon sales and mitigation levels.

While carbon trade implements any abatement target at the lowest cost, it has proven to be difficult to establish consensus on binding targets. Helm's (2003) permit trade game with endogenous permit choice responds to this difficulty. While one is tempted to think that endogenous permits choice makes it attractive for all players to join the market, Helm (2003) shows that this is not true. In a global market with endogenous permit choice high damage countries may suffer a welfare loss compared to a Nash equilibrium in the emissions game without trading. This suggests that a partial market can be established by a subgroup of countries. Carbone et al. (2009) analyse incentives to join an international carbon market with the application of Helm's idea of endogenous allowance choices to a calibrated general equilibrium model. Their numerical results show that sub-global carbon markets exist and these equilibrium markets are successful to induce the participation of developing countries. In Carbone et al. (2009), the results of stable coalitions are obtained only by considering internal stability of the carbon market. In this paper, we apply the solution concept of the cartel stability where the carbon market is stable when both of internal and external stability is satisfied. Our results show that only a partial carbon market emerges and a mitigation coalition would usually not join the market. The intuition is that the mitigation coalition countries would usually have high (aggregate) marginal abatement benefits, they would be permit buyers thereby the negative transfers of buying emission permits and the benefits from free-riding surpass the gain from trading.

Since the aim of the study is to go beyond qualitative results and to obtain insights in the relative strength of the different interacting effects of a mitigation agreement and carbon trade with partial agreements and partial markets we employ an adjusted version of the STAbility of COalitions model (STACO 3.0). Our version of STACO specifies the abatement benefit and cost functions for seven regions to evaluate the effects of endogenous permits choice on the stability of international coalitions and abatement efficiency. The original STACO 3.0 has been developed by Nagashima et al. (2011) and Dellink et al. (2015) and comprises 12 regions.

The structure of this paper is as follows. In the next section, we introduce our four-stage game. In Section 2.3, we present the analytical results. In section 2.4, we provide a brief description of STACO 3.0 and explain our re-calibration. In Section 2.5, we present numerical results on the formation and efficiency of a mitigation agreement and a carbon market with endogenous permits choice. Section 2.6 concludes.

2.2 The model

We consider a game with a set of countries or regions $N = \{1, \dots, n\}$ that aim at maximising individual payoffs in a climate policy game. Our game has four stages: (i) formation of a mitigation agreement, (ii) formation of a carbon market, (iii) choice of emission permits, and (iv) carbon trade and abatement. A detailed description of the four stages follows.

Stage 1: Formation of a Mitigation Agreement

At the first stage, each player $j \in N = \{1, \dots, n\}$ announces whether or not to join a mitigation agreement. Formally each player has a binary strategy space {join, not join}. We refer to the subset of players who join as the signatories or the mitigation coalition $M \subseteq N$; we refer to the non-signatories as singletons. M is called "grand coalition" if $M = N$. If $M = \{j\}$ or $M = \{\}$, we have the All-singletons structure. M is a non-trivial partial coalition if it has at least two members. The remaining singletons are represented as the set $N - M$.

We assume that signatories coordinate their climate policies to maximise joint payoffs. Hence, at subsequent stages, the coalition M acts as a single player. Our game has three subsequent stages and we employ the concept of Partial Agreement (Subgame Perfect) Nash Equilibrium (Chander & Tulkens 1995 and 1997) to solve our game. We will say that a mitigation coalition is stable if no signatory has an incentive to leave and no non-signatory has an incentive to join.

Stage 2: Formation of a Carbon Market

At this stage, the coalition M acts jointly and the remaining singletons act individually, to decide to join or not to join the market for the trade in carbon emission permits. Formally each player has, again, a binary strategy space {join, not join}. We refer to the subset of players who join the market as traders $T \subseteq N$; Clearly, in order to be effective, T must have at least two members. The carbon market, if it is formed, is assumed to be competitive and to satisfy Walras' law. We will say that market participation is stable if no trader has an incentive to leave and no non-trader has an incentive to join the market.

Stage 3: Choice of Initial Permits Endowment

At this stage, only market participants denoted by $k \in T$ make a choice. They choose an amount of initial emission permits ω_k. We refer to these as market participant k's endowment. Coalition members jointly announce their aggregate endowment. The aggregate of all announcements $\omega = \sum_{k \in T} \omega_k$ determines the size of the market and the aggregate abatement level of all trade participants. At this stage, the carbon price denoted by p (=$p(\omega)$) is determined by the total number of permits and equals to marginal abatement costs in equilibrium. Hence, with the assumption of market clearing individual trader's abatement level decided at stage 4 can be indirectly affected by the total permit choice $q_k = q_k(p(\omega))$. The choice of each market participant depends on three factors: benefits from global abatement denoted by B_k, costs from own abatement denoted by C_k and revenues from trade. In the model of this paper, players' abatement benefits B_j, which depends on the global abatement level $q = \sum_{j \in N} q_j$, are defined as linear form ($B_j'(q) > 0, B_j''(q) = 0$). This linear form assures the dominant strategy of all non-participants' abatement choice. Players' mitigation costs C_j depend on a country's own mitigation effort q_j and are convex ($C_j'(q_j) > 0, C_j''(q_j) > 0$).

Stage 4: Permits Trade and Abatement

At the final stage all players choose their mitigation levels. It is assumed in the model that each player's abatement level does not exceed uncontrolled emissions denoted by $\bar{e}_j$. At this stage we have four possible types of players: singleton non-traders, singleton traders, members of a non-trading coalition, and members of a trading coalition.

The general structure of the four-stage game is summarized in Figure 2.1:

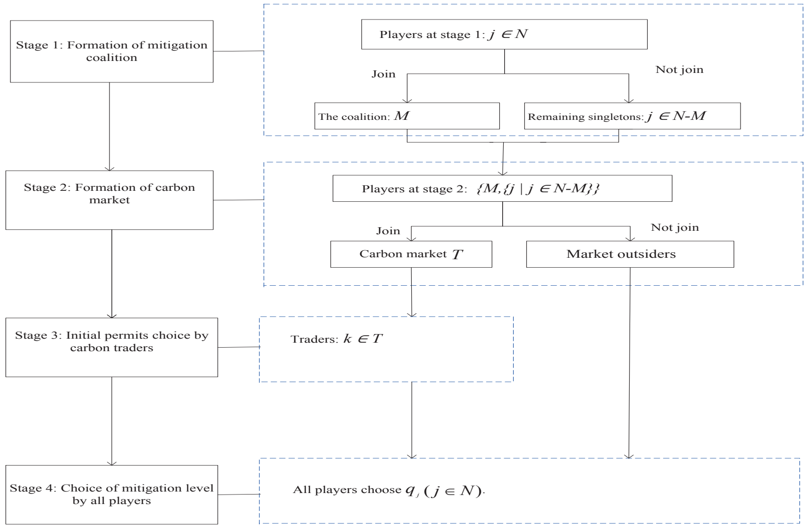

Figure 2.1. Schematic representation of the game proceeding with four stages

2.3 Analysis

In this section, we present the theoretical analysis and results for each stage. We apply Subgame Perfect Nash Equilibrium and solve the game by backward induction.

Stage 4: Mitigation

At this stage, each player has decided whether or not to be a signatory to the mitigation agreement and whether or not to participate in the carbon market. Market participants have announced their endowment such that $\omega \equiv \sum_{k\in T} \omega_k$ and (implicit) permit price is given at this stage. Best response abatements will depend on the choices made at previous stages. Thus we identify conditions for the best responses for each of the four types of players.

a. For singleton non-traders $j \notin T, j \notin M$, they choose their optimal mitigation level by solving the following problem:

$$\max_{q_j} \ [\pi_j = B_j(q) - C_j(q_j)]. \tag{2.1}$$

The first order condition is obtained as

$$B_j'(q^*) = C_j'(q_j^*). \tag{2.2}$$

Singleton non-traders just equate their marginal benefits with their marginal costs of mitigation.

b. A singleton trader $k \in T, k \notin M$ must set its mitigation level q_k according to the number of permits it holds after buying or selling permits. The optimization problem is:

$$\max_{q_k} \ [\pi_k = B_k(q) - C_k(q_k) + p[\omega_k - (\bar{e}_k - q_k)]. \qquad (2.3)$$

The third term of right side of Eq. (2.3) shows trade revenues, and the quantity traded is the difference between permits ω_k and after-trade emissions $e_k(= \bar{e}_k - q_k)$. The best response by singleton traders is derived by taking the derivative of Eq. (2.3) with respect to q_k:

$$p^* = C_k'(q_k^*) \, . \qquad (2.4)$$

For traders it is optimal to equate marginal mitigation costs with the permit price.

c. If the coalition participates in the carbon market, then the coalition member that also belongs to the carbon market $k \in M \cap T$ jointly solves

$$\max_{q_k} \ [\sum_{i \in M} \pi_i = \sum_{i \in M} (B_i(q) - C_i(q_i) + p[\omega_i - (\bar{e}_i - q_i)])]. \qquad (2.5)$$

The best response of a coalition member when the coalition is participating in carbon trade must also satisfy condition (2.4). This follows from the first order condition obtained from problem (2.5). Recall that at this stage members of a mitigation agreement coordinate their climate policies and the coalition acts as a single player, i.e. all or none of the coalition members enter the carbon market. In a competitive market, trade equalises the marginal abatement costs among traders through a uniform permits price p^*. For a convex cost function (2.4) implicitly gives individual abatement as a function of price. Since the carbon market must clear, we also have $\omega = \sum_{k \in T}(\bar{e}_k - q_k)$ and we can determine the equilibrium price.

d. If the coalition does not participate in the market, then members $j \in M, j \notin T$ solve the following problem:

$$\max_{q_j} \ [\sum_{i \in M} \pi_i = \sum_{i \in M} \big(B_i(q) - C_i(q_i)\big)]. \qquad (2.6)$$

The first order condition is:

$$C_j'\big(q_j^*\big) = \sum_{i \in M} B_i'(q^*). \qquad (2.7)$$

Joint payoff maximisation requires that coalition members equate individual marginal cost with the sum of all members' marginal benefits (Samuelson's rule).

With these results we can move to the analysis of stage 3 of the game.

Stage 3: Choice of Initial Permit Endowments

At this stage only traders make a choice. A singleton trader's best response initial permits choice ω_k^* is the solution to the following problem:

$$\max_{\omega_k} \ [\pi_k = B_k\big(q(\omega)\big) - C_k\left(q_k^*\big(p^*(\omega)\big)\right) + p^*(\omega)\left(\omega_k - \bar{e}_k + q_k^*\big(p^*(\omega)\big)\right)], \ \ k \in T, k \notin M. \tag{2.8}$$

For convenience we will sometimes refer to emissions instead of abatement:

$$e_k^*\big(p^*(\omega)\big) \equiv \bar{e}_k - q_k^*\big(p^*(\omega)\big).$$

Differentiating the objective function Eq. (2.8) with respect to ω_k yields the equilibrium outcome:

$$\omega_k^* = e_k^* + \frac{B_k'\left(q^*\big(p^*(\omega^*)\big)\right) - C_k'\left(q_k^*\big(p^*(\omega^*)\big)\right)}{p^{*\prime}(\omega^*)}. \tag{2.9}$$

If the coalition participates in the carbon market the best response choice of initial permits ω_k^* solves the following problem of joint payoff maximisation:

$$\max_{\omega_k} \ [\sum_{i \in M} \pi_i(\omega_i) = \sum_{i \in M} \left[B_i\left(q^*\big(p^*(\omega)\big)\right) - C_i\left(q_i^*\big(p^*(\omega)\big)\right)\right] + p^*(\omega) \sum_{i \in M}\left(\omega_i - e_i^*\big(p^*(\omega)\big)\right)], k \in M \cap T. \tag{2.10}$$

Taking derivatives of Eq. (2.10) with respect to ω_k, we obtain from the first order condition

$$\sum_{i \in M} \omega_i^* = \sum_{i \in M} e_i + \frac{\sum_{i \in M} B_i'\big(q^*(\omega^*)\big) - C_j'\left(q_j^*\big(p^*(\omega^*)\big)\right)}{p^{*\prime}(\omega^*)}. \tag{2.11}$$

The final step in the analysis of the stage-3 game is to sum up the equilibrium announcements of all market participants $k \in T$ using Eq. (2.9) and Eq. (2.11). Using the market clearance condition and individually optimal abatement $p^* = C_k'(q_k^*)$ (see Eq. (2.4)) we obtain the equilibrium price of permits:

$$p^*\big(q^*(\omega^*)\big) = \frac{1}{\# T}\sum_{k\in T} B_k'\big(q^*(\omega^*)\big). \tag{2.12}$$

where $\# T$ represents the number of market participants and is specified as follows:

$$\# T \equiv \begin{cases} |(N-M)\cap T|, & \textit{if } M \textit{ does not join the market} \\ |(N-M)\cap T|+1, & \textit{if } M \textit{ joins the market}. \end{cases}$$

Notice that the mitigation coalition acts as a single player. From Eq. (2.12), we can see that the equilibrium price depends on the number of the market participants and the marginal abatement benefits of all traders.

Furthermore, in line with Helm (2003), we obtain the following conclusion from Eqs. (2.9) and (2.11). A singleton is a permit buyer (seller) if its marginal benefits of abatement are larger (smaller) than its marginal costs of abatement. Formally it follows from (2.9) that

$$\begin{cases} \textit{permits buyer}: e_k^* - \omega_k^* > 0 \Leftrightarrow C_k'(q_k^*) < B_k'(q^*) \\ \textit{permits seller}: e_k^* - \omega_k^* < 0 \Leftrightarrow C_k'(q_k^*) > B_k'(q^*) \end{cases}, k \in (N-M)\cap T. \tag{2.13}$$

A coalition member is a permit buyer (seller) if the sum of the coalition's marginal benefits of abatement are larger (smaller) than that member's marginal costs of abatement. Formally it follows from Eq. (2.11) that

$$\begin{cases} \textit{permits buyer}: \sum_{i\in M} e_i^* - \sum_{i\in M}\omega_i^* > 0 \Leftrightarrow C_k'(q_k^*) < \sum_{i\in M} B_i'(q^*) \\ \textit{permits seller}: \sum_{i\in M} e_i^* - \sum_{i\in M}\omega_i^* < 0 \Leftrightarrow C_k'(q_k^*) > \sum_{i\in M} B_i'(q^*) \end{cases}, k \in M\cap T. \tag{2.14}$$

This result means that we can identify buyers and sellers by looking at the relationship between marginal costs and marginal benefits of abatement. Helm (2003) therefore distinguishes low-damage ($B_k'(q^*) < C_k'(q_k^*)$) and high-damage ($B_k'(q^*) > C_k'(q_k^*)$) countries, where in our case the damages are the avoided benefits from abatement. It is clear from equations (2.13) and (2.14) that a coalition member is more likely to be a permit buyer as the sum of marginal benefits for the coalition is more likely to exceed individual marginal costs.

Stage 2: Market Participation Choice

Now we move back to stage 2. The sub-game played at this stage determines trade participation. During this stage, players, i.e. the coalition M and the singletons, simultaneously choose to join or not to join T. The choice of market participation depends on the payoffs from being a carbon trader or an outsider. The analysis of stages 3 and 4 of the game implicitly defines payoffs for each coalition and the remaining singletons for every possible partition of the players into a set of traders and outsiders. In other words: payoffs can be written as a function of the coalition structure and market participation. We denote the resulting partition function by $V_j(M,T)$. At stage 2, given M, player j would participate in the market T if and only if

$$V_j(M,T) \geq V_j(M,T\backslash\{j\}). \tag{2.15}$$

In the equilibrium, condition (2.15) must hold for traders. In addition, outsiders must prefer not to participate in trade:

$$V_j(M,T) \geq V_j(M,T \cup \{j\}). \tag{2.16}$$

Conditions (2.15) and (2.16) are the internal and external stability conditions for market participation.

Stage 1: Formation of a mitigation coalition

Moving back to the first stage, i.e. the formation of mitigation coalition M, each player $j \in N$ makes choice of joining or not joining coalition M. The choice of membership of M depends on the payoffs from being a coalition member or a singleton. The internal and external stability conditions of the mitigation coalition is:

$$V_j(M) \geq V_j(M\backslash\{j\}), j \in M, \tag{2.17}$$
$$V_j(M) \geq V_j(M \cup \{j\}), j \notin M. \tag{2.18}$$

Within M, each signatory is confronted with the incentive to be a free-rider. To alleviate this effect, we apply a transfer rule from the class of optimal transfers (Carraro et al. 2006, Weikard 2009) that allocates the coalition payoff proportional to outside option payoffs:

$$\mathcal{V}_j(M,T) = \frac{V_j(M\backslash\{j\})}{\sum_{i\in M} V_i(M\backslash\{i\})} V_M(M), j \in M. \tag{2.19}$$

where we denote aggregated payoffs of coalition M by $V_M(M)$. $V_j(M\backslash\{j\})$ is the outside payoff of coalition member $j \in M$. We use $\mathcal{V}_j(M)$ to denote payoffs after transfers.

2.4 Numerical analysis based on STACO model

2.4.1 Background of STACO 3.0

To gain a more explicit and practical insight in our game we apply an integrated assessment model STACO 3.0 (Nagashima et al., 2011; Dellink et al., 2015) for a numerical analysis. STACO is an integrated model which links a (simple) climate change module with game theory for economic and policy analysis. The climate module of STACO relates GHG emission paths to GHG concentrations and atmospheric temperature change. The model has been calibrated in line with the EPPA model (Paltsev et al., 2005) and the DICE model (Nordhaus, 1994; Nordhaus, 2008). The economic part of STACO specifies regional payoff functions for emissions mitigation, which is composed of abatement costs and benefits. Regional abatement benefits, which are calculated as the reduced damages, are given as a share of global abatement benefits (Fankhauser, 1995; Tol, 1997; Tol, 2009). STACO formulates a two-stage game of ICA formation among 12 heterogeneous regions: United States (USA), Japan (JPN), European Union-27 & EFTA (EUR), Other High Income countries (OHI), Rest of Europe (ROE), Russia (RUS), High Income Asia countries (HIA), China (CHN), India (IND) and the Middle East countries (MES), Brazil (BRA) and Rest of the World (ROW).

The main equations of STACO 3.0 are presented in Box 2.1. STACO adopts a time horizon of 100 years, ranging from 2011 ($t = 1$) to 2110 ($t = 100$). Eqs. (2.20) to (2.22) show the objective functions for coalition members $j \in M$, market participants $k \in T$ and singletons $j \notin M$, $j \notin T$, which are based on the net present value of payoffs accruing to regions over a period of 100 years. Eq. (2.23) shows a linear functional form of abatement benefits, where s_j represents the regional share of global benefits with $\sum_{j \in N} s_j = 1$, and γ_D gives the aggregated climate change damages in terms of a percentage of gross world product (GWP). Parameter $\delta_{j,\tau}$ reflects two impacts of mitigation adopted in current period t: one is the impact on future climate due to the inertia in the climate system; the other is the impact on GDP growth. The calibration of $\delta_{j,\tau}$ is based on the EPPA-5 (Paltsev *et al.,* 2005) model by using a climate module from the DICE (Nordhaus, 1994) model. The regional abatement cost function (Eq. (2.24)) is a cubic function of regional abatement efforts $q_{j,t}$, where parameters α_j and β_j are parameters estimated based on the data from EPPA (Morris *et al.,* 2008), which is shown in Table 2.A1. For discount rates $r_{j,t}$, STACO 3.0 uses the Ramsey rule, which implies a changing discount rate over time based on the pure rate of time preference and regional growth rates of GDP. STACO 3.0 uses data from EPPA-5 model to calibrate the regional business as usual (BAU) emission paths (see Table 2.A1 for the BAU emissions in period one).

Box 2.1. Main model equations of STACO 3.0

Payoff functions (Objective functions)

$$\max_{q_{j,t}} \pi_{j,t}(q_t) = \sum_{i\in M}\sum_{t=1}^{t=100}\{(1+r_{i,t})^{-t}\cdot(B_{i,t}(q_t)-C_{i,t}(q_{i,t}))\}, \qquad \forall j \in M \qquad (2.20)$$

$$\max_{q_{k,t}} \pi_{k,t}(q_t) = \sum_{t=1}^{t=100}\{(1+r_{k,t})^{-t}\cdot(B_{k,t}(q_t)-C_{k,t}(q_{k,t})+p_t[\omega_{k,t}-(\bar{e}_{k,t}-q_{k,t})])\}, \forall k \in T \qquad (2.21)$$

$$\max_{q_{j,t}} \pi_{j,t}(q_t) = \sum_{t=1}^{t=100}\{(1+r_{j,t})^{-t}\cdot(B_{j,t}(q_t)-C_{j,t}(q_{j,t}))\}, \qquad \forall j \notin M, j \notin T \qquad (2.22)$$

with global abatement: $q_t = \sum_{j\in N} q_{j,t}$.

Abatement benefits

$$B_{j,t}(q_t) = \sum_{\tau=t}^{\infty}\left\{(1+r_{j,\tau})^{t-\tau}\cdot(s_j\cdot\gamma_D\cdot\delta_{j,\tau}\cdot q_t)\right\}, \qquad \forall j \in N \quad (2.23)$$

which captures the long term effect of current abatement $q_{j,t}$ *on climate change by including future time periods* $\tau \subseteq [t,\infty)$.

Abatement costs

$$C_{j,t}(q_{j,t}) = \left(\frac{1}{3}\cdot\alpha_j\cdot q_{j,t}^3 + \frac{1}{2}\cdot\beta_j\cdot q_{j,t}^2\right)\cdot\chi_{j,t}, \qquad \forall j \in N \qquad (2.24)$$

where $0 < \chi_{j,t} < 1$ *is declining over time to reflect cost savings due to technological progress.*

Ramsey rule for the discount rate

$$r_{j,t} = \rho + \eta\cdot\left(\frac{GDP_{j,t}}{GDP_{j,t-1}} - 1\right), \quad \rho = 0.015, \qquad \eta = 1. \qquad \forall j \in N \qquad (2.25)$$

2.4.2 Modified STACO 3.0: aggregation of regions

The software package MATLAB (R2012a) is used for the numerical analysis. The partition function for the12-region STACO 3.0 model for mitigation coalitions is defined on the power set of the set of players N. In our case, since we examine coalition membership and trade participation at the same time, the partition function is defined on the power set of $N \times N$. For a numerical model with 12 regions the partition function cannot be calculated within reasonable time. We therefore aggregate the 12 regions into 7 regions: United States (USA), European Union-27 & EFTA (EUR), China (CHN), India (IND), the region of High Abatement costs and High Income (HAHI), [1] the region of Low Abatement costs and Energy Exporting (LAEX) [2] and Rest of the World (ROW). This reduces the calculation time by three orders of magnitude and allows for the numerical calculation of all possible coalitions and their payoffs in our model. We have recalibrated the STACO parameters for the abatement cost functions of the seven regions, and the undiscounted marginal cost curves in base year 2011 are shown in Figure 2.2:

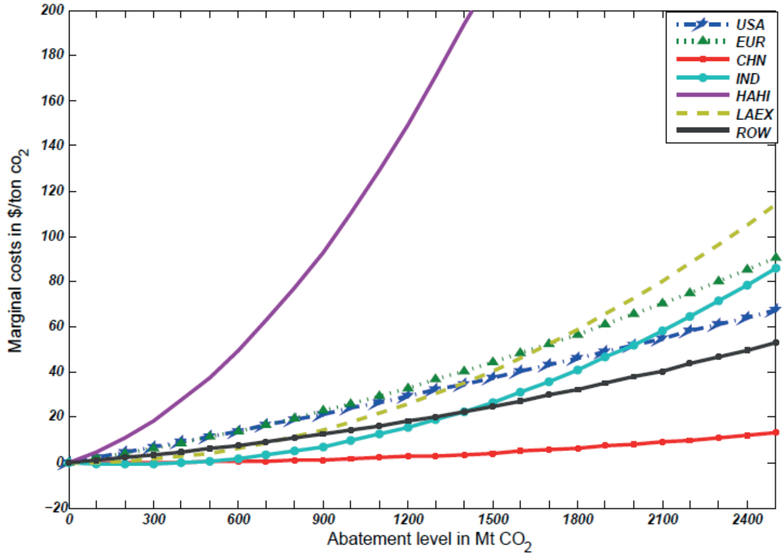

Figure 2.2. Marginal abatement cost curves in 2011 for STACO model with 7 regions

All other parameters for STACO 3.0 with 7 regions are shown in Table 2.A1 in 2.7 Appendix II.

2.4.3 Implications from the specified functions

Based on the functions specified in Box 1, the relationship between equilibrium permit price p_t^* and the total number of optimal initial permits ω_t^* is (the derivation is provided in 2.7 Appendix II):

$$\omega_t^* = \bar{e}_t - \sum_{k\in T} \frac{-\beta_k \cdot \chi_{k,t} + \sqrt{(\beta_k \cdot \chi_{k,t})^2 + 4\alpha_k \cdot \chi_{k,t} \cdot p_t^*}}{2\alpha_k \cdot \chi_{k,t}}. \tag{2.26}$$

[1] The integration of four regions in original STACO model: JPN, OHI, BRA and HIA.
[2] The integration of three regions in original STACO model: ROE, RUS and MES.

Eq. (2.26) cannot be solved analytically in an explicit form of $p_t^*(\omega_t^*)$, therefore we resort to a numerical analysis. If we assume, for example, the number of traders is 7 in period $t = 1$, then the numerical value of p_t^* can be obtained for any value ω_t^*. Figure 2.3 shows a plot of the relationship between p_t^* and ω_t^* in the carbon market. As expected, the price of permits falls when the supply of permits increases, and the price decreases to zero when the total initial permits are chosen as the BAU emissions level.

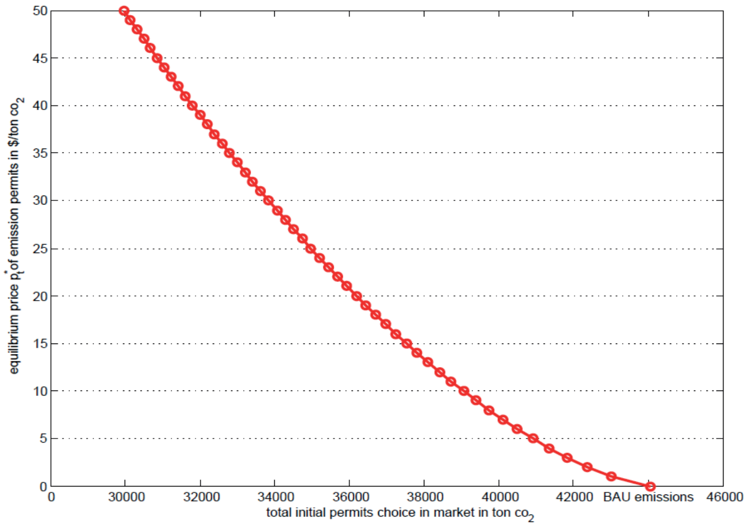

Figure 2.3. Relationship between equilibrium price and total number of initial permits in carbon market

By taking the derivative of both sides of Eq. (2.26) with respect to ω_t^*, we get a differential equation of permit price p_t^*:

$$p_t^{*\prime}(\omega_t^*) = -\frac{1}{\sum_{k\in T}\left(\left(\beta_k\chi_{k,t}\right)^2 + 4\alpha_k\chi_{k,t}p_t^*\right)^{-\frac{1}{2}}}. \tag{2.27}$$

Note that the coalition enters the carbon market as a single agent and is included in T. From Eq. (2.27), we can see that under higher permit prices, the marginal price due to one more unit change in permits supplying (the absolute value $\left|p_t^{*\prime}(\omega_t^*)\right|$) is higher than the one under lower prices. This also can be observed from Figure 2.3.

2.5 Results

In this section we show the results in three parts. Firstly, in subsection 2.5.1, we analyse the benchmark scenario with a single mitigation agreement M not considering carbon trade, which includes two cases: All-Singletons $M = \{j\}$ or $M = \{\}$, and the Grand Coalition $M = N$; then, in subsection 2.5.2, we analyse stability and efficiency of partial mitigation agreement M without a carbon market; in subsection 2.5.3, we introduce carbon trade with endogenous permits choice in this setting.

2.5.1 Two cases for the benchmark scenario: All-Singletons and Grand Coalition without carbon trade

We provide two cases as benchmark scenarios: All-Singletons (AS) and Grand Coalition (GC). In All-Singletons, each region chooses its own mitigation as a best response to others' mitigation efforts ($C_j'(q_j) = B_j'(q)$, $j \in \mathrm{N}$, see Table 2.1, columns 6 and 7), which corresponds to the non-cooperative Nash equilibrium of the abatement game. The Grand Coalition is the best-performing coalition in terms of global abatement and net benefits and implements full cooperation. Under the GC, the marginal abatement costs of all regions are equalized with the sum of marginal benefits, i.e. $C_j'(q_j) = \sum_{j \in \mathrm{N}} B_j'(q)$.

Comparing the two cases, we find that: firstly, the abatement level (see columns 2 and 3 in Table 2.1) by each GC member increases as compared to the level in AS, which is reflected by the increased marginal abatement costs (see columns 8 and 9 in Table 2.1). The three regions with the largest increase in abatement levels are CHN, IND and ROW. This is due to their relatively low marginal abatement costs (see Figure 2.2) which causes them to take larger shares of the global abatement. The same happens to the payoffs presented in columns 4 and 5 of Table 2.1, where USA, EUR and HAHI take the largest shares of the total benefits because they have the highest marginal benefits (see column 6 in Table 2.1) and lower mitigation burdens (see column 3 in Table 2.1). Secondly, the optimal transfer scheme applied in our model is used to alleviate the imbalance between mitigation burdens and net benefits, which implies a flow of transfers from regions with higher marginal benefits, e.g. USA, EUR, HAHI and LAEX, to regions with large mitigation burdens, e.g. CHN, IND and ROW (see column 11 in Table 2.1). Note that Ramsey's rule (Eq. (2.25)) is used for discounting, so that different discount rates are applied for different regions. With respect to the internal stability, we find that each GC member prefers to leave the coalition (internal instability), because the gains from cooperation are not sufficient to compensate for all players' free-riding incentives.

The first best mitigation policy of the GC offers a large increase in global mitigation and offers the largest possible welfare gains. However, every region has a strong incentive to free-ride, especially the regions with high marginal abatement benefits and costs (e.g. USA, EUR and HAHI).

Table 2.1. All Singletons (AS) and Grand Coalition (GC) in a setting without a carbon market

	Mitigation in 2011 in % of BAU emissions		NPV of payoffs over 100 years (billion $)		Marginal abatement benefits ($/ton CO_2)		Marginal abatement Costs ($/ton CO_2)		Incentive to leave GC (billion $)	NPV of Transfers within GC (billion $)*
	AS	GC	AS	GC	AS	GC	AS	GC	GC	GC
USA	8.5	33.0	6347.9	11192.4	14.1	14.1	14.1	62.3	3799.5	-2416.9
EUR	16.0	39.4	8928.4	14944.8	19.3	19.3	19.3	62.3	5430.6	-7065.0
CHN	7.7	45.4	519.0	2935.2	1.2	1.2	1.2	62.3	284.5	3470.5
IND	15.5	66.1	329.4	2554.9	0.7	0.7	0.7	62.3	208.7	927.0
HAHI	3.8	9.3	8304.1	13210.6	17.6	17.6	17.6	62.3	5051.9	-8074.2
LAEX	12.6	39.1	2842.5	8931.9	6.1	6.1	6.1	62.3	1842.7	-119.3
ROW	5.1	49.1	1482.0	6860.7	3.2	3.2	3.2	62.3	923.3	2785.8
World	8.8	38.0	28753.3	60630.4	–	–	–	–	–	–

* Since regional discount rates are used, transfers do not add up to zero.

2.5.2 Mitigation coalition M without a carbon market

In this section, we investigate the formation of stable coalitions and their mitigation efficiency without considering the carbon market. This game is a standard two-stage coalition formation game. We check the stability of all 127 (i.e. $2^7 - 1$) mitigation coalitions under the optimal sharing rule, and find that 28 non-trivial coalitions are stable. The results are shown in Table 2.A2 in 2.7 Appendix II, where all stable coalitions are ranked by an "indicator of success" [3] (last column in Table 2.A2), which reflects the relative welfare improvement by stable coalitions. The two best performing coalitions (in terms of welfare gains from cooperation) comprise the largest number of signatories as compared to the other coalitions (see column 1 in Table 2.A2). These results reflect the following. Firstly, a stable coalition is more easily formed by a small set of signatories, since larger coalitions are associated with stronger free-riding incentives, which causes deviation from cooperation. Secondly, the minimum number of signatories in all stable coalitions is no less than 3, which means that all two-member coalitions are internally stable. This finding reflects the conclusion of Weikard et al. (2006) that all two-player coalitions are internally stable under a linear abatement benefit function and non-negative claims with proportional surplus sharing. Thirdly, the results show the superiority of the larger coalitions in terms of mitigation efficiency, that is, coalitions with

[3] Indicator of success (%) is defined as [(NPV of global payoff in a coalition-NPV of global payoff in All Singletons)/(NPV of global payoff in Grand coalition-NPV of global payoff in All Singletons)]*100.

more members have a higher indicator of success. Furthermore, from the coalitional structure, we can observe that all stable coalitions are combinations of regions with higher marginal abatement costs and benefits on the one hand and regions with lower marginal abatement costs and benefits on the other, e.g. USA-CHN (or USA-IND), EUR-CHN (or EUR-IND) and HAHI-CHN (or HAHI-IND). This structure counterbalances the free-riding incentives for signatories with relatively high abatement costs through higher mitigation levels by signatories with low abatement costs. The compensation is transferred from high benefit gaining signatories to losing signatories. Comparing coalitions with higher and lower indicators of success, generally coalitions with lower indicators comprise more signatories with expensive abatement than signatories with cheap abatement options. Example are the differences between the first two coalitions ({CHN, IND, HAHI, ROW} and {EUR, CHN, IND, ROW}) and the last two coalitions ({USA, HAHI, LAEX} and {EUR, HAHI, LAEX}) (see Table 2.A2). This shows that the presence of regions with low marginal costs improves the efficiency of a coalition in terms of the net abatement benefits.

Inspection of the structure of high ranking stable coalitions shows that participation by CHN reduces global abatement costs and improves global welfare. China's participation seems to be crucial. All stable coalitions where China participates show a larger indicator of success than any of the stable coalition (see Table 2.A2). This finding shows that it is essential to engage China in GHG emissions mitigation policies, which is consistent with the result of Paltsev et al. (2012) who analysed the role of China in mitigating greenhouse gas emissions. Any one of the top 3 regions in terms of the highest marginal abatement costs (see Figure 2.2), HAHI, LAEX or EUR, can form a stable coalition with any other region with low abatement costs such as CHN, IND and ROW.

Table 2.2 shows details for the stable coalition with the highest global welfare {CHN, IND, HAHI, ROW}. Within this coalition, members equalize their marginal abatement costs, which results in large gains of 11685.7 billion \$ for HAHI, but HAHI compensates CHN, IND and ROW by 4386.4 billion \$. Under this structure, all outsiders can benefit from the positive externality of increased global abatement, which causes the negative incentives to change membership for all regions.

Table 2.2. The best-performing M:{CHN, IND, HAHI, ROW}

	Mitigation in 2011 in % of BAU emissions	NPV of payoffs over 100 years (billion US $)	Marginal abatement Costs ($/ton CO_2)	Incentive to change membership (billion $)	NPV of transfers within M (billion $)*
USA	8.5	12835.3	14.1	-1300.2	–
EUR	16.0	17935.0	19.3	-1793.5	–
CHN	28.0	753.1	22.7	-12.3	853.1
IND	42.9	564.8	22.7	-9.0	256.7
HAHI	4.7	11685.7	22.7	-183.7	-4386.4
LAEX	12.6	5622.8	6.1	-443.6	–
ROW	24.9	2495.1	22.7	-44.1	683.2
World	18.7	51891.8	–	–	–

* Since regional discount rates are used, transfers do not add up to zero.

2.5.3 Mitigation coalition M under a carbon market with endogenous permits choice

In this section, our analysis focuses on the impact of carbon trade on the stability and efficiency of the mitigation coalition M. For brevity of notation, we use the set structure $\{M, T\}$ to represent each coalition M under a particular market structure T. Firstly, we calculated the abatement level and payoffs for each region under different $\{M, T\}$ structures, and then we analysed the internal and external stability of each structure. Table 2.3 shows the results: there are 3 stable $\{M, T\}$ structures among the total of 3990 structures, ranked by their indicators of success. The second column of Table 2.3 lists the stable coalition M, and the subsequently formed T are presented in the third column. The three stable coalitions all contain the same two regions IND and ROW which cooperate respectively with USA, HAHI and EUR on mitigation. In equilibrium the choice of coalition membership is combined with the choice of market participation by CHN and LAEX in each of the three structures $\{M, T\}$. Here CHN takes the role of a permit seller in the carbon market because it has the cheapest abatement options. One interesting point we find is that for all three stable structures $\{M, T\}$, the abatement coalition M has no incentive to join the carbon market T which is composed of non-signatories only. Our numerical results also show that if M would join the carbon market, then M would always be a buyer, implying negative transfers for M. Furthermore, M can benefit from the positive externality of increased mitigation efforts resulting from a carbon market formed by singletons. The best-performing structure is $\{M$={USA, IND, ROW}, T={CHN, LAEX}$\}$, which corresponds to the highest value of success indicator and also reaches the highest abatement level, but shows a significantly lower global payoffs than the best-performing coalition in the absence of a carbon market, shown in Table 2.2.

Figure 2.4 shows global abatement levels over time under the three $\{M, T\}$ structures for the time horizon of 100 years. Overall, the abatement commitments of three coalition structures increase over time. Due to the identical markets across the three stable structures with the same mitigation level of traders, differences in mitigation over structures are exclusively driven by the mitigation efforts of the respective coalitions. More specifically, the level of abatement is a function [4] of coalitional marginal benefits $\sum_{i \in M} B'_{i,t}$ which vary across coalitions M. As can be seen from Figure 2.4, the global abatement of structure 3 takes the lead initially due to the ranking of $B'_{EUR} > B'_{HAHI} > B'_{USA}$. While this ranking changes after 35 years, the global abatement level of structure 1 surpasses that of structure 3 and the disparity increases after that. This can be explained by i) the decreased gaps of marginal benefits between EUR and USA, which reduce the difference between the total abatement levels of $M3$ and $M1$, and by ii) the value of parameters α_{EUR} and α_{USA} in the abatement cost function. The same holds for coalition structure 2 and structure 3 from the 85th year onwards, the global abatement levels of structure 2 exceed those of structure 3, which is due to the marginal abatement costs of HAHI, which are larger than those of EUR from the 85th year onwards.

Table 2.3. The stable structures $\{M, T\}$

M		T		NPV of global payoffs over 100 years (billion \$)	Global Mitigation in full-period in % of BAU emissions	Indicator of success
M1	USA, IND, ROW	T1	CHN, LAEX	40160.3	23.7	35.8
M2	IND, HAHI, ROW	T2	CHN, LAEX	39533.5	23.4	33.8
M3	EUR, IND, ROW	T3	CHN, LAEX	39506.3	23.5	33.7

[4] Based on the Eq.(2.11) in Section 2.3 and the specified model equations (shown in Box 2.1) in STACO 3.0, the abatement level of coalition member $k \in M$ at steady state is $q^*_{k,t} = \frac{-\beta_k \cdot \chi_{k,t} + \sqrt{(\beta_k \cdot \chi_{k,t})^2 + 4\alpha_k \cdot \chi_{k,t} \cdot \sum_{i \in M} B'_{i,t}}}{2\alpha_k \cdot \chi_{k,t}}$, where $\sum_{i \in M} B'_{i,t}$ is the coalitional marginal benefits.

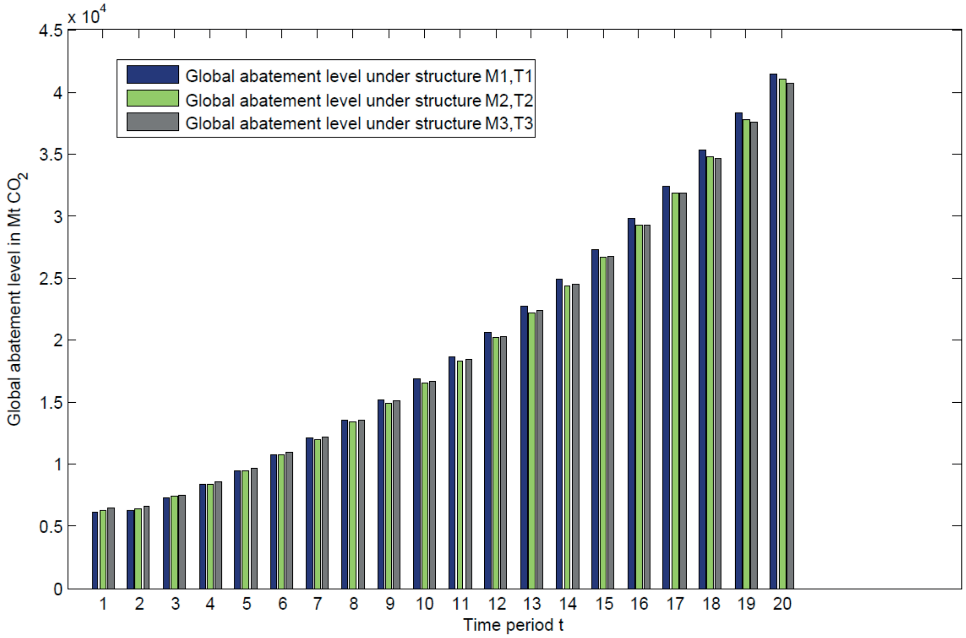

Figure 2.4 Global abatement path under three stable structures

As our particular interest is in the impacts of the carbon market on the stability and efficiency of mitigation coalitions, the comparison between a stable mitigation coalition without the carbon market and the coalition with carbon market $\{M, T\}$ is displayed in Tables 2.4, 2.5a and 2.5b.

First, the coalition M={USA, IND, ROW}, which is stable without a carbon market, is also stable under a carbon market with the structure of $\{M$={USA, IND, ROW}, T={CHN, LAEX}$\}$. From Table 2.4, it can be observed that global mitigation levels and global welfare are both improved under T. The indicator of success for coalition M={USA, IND, ROW} (see the last column in Table 2.3) also increases as compared to the indicator under a single coalition M (see the last column in Table 2.A2). Furthermore, in coalition M, the abatement level of each coalition member remains the same. The payoffs, however, increase when CHN and LAEX create a carbon market. This improvement is driven by the positive externality of the increase in the global mitigation level arising from the carbon market. The global mitigation level increases from 5790.6 Mton CO_2 to 6177.3 Mton CO_2 with the presence of the carbon market. The payoffs of both market participants CHN and LAEX increase due to the relatively lower carbon price and the large positive surplus of payoffs through trading. Due to lower mitigation costs, CHN prefers to export emission permits, while LAEX, which has relatively higher abatement costs, prefers to buy permits in the market. As a result, CHN achieves a higher abatement level and gains revenues from selling, while LAEX abates less as a permits buyer than as a market outsider. Outsiders of both M and T, EUR and HIHA, have the same mitigation commitments in these two scenarios, however, their payoffs are improved due to the increased global abatement with a carbon market. In this case, with a given stable coalition M, the presence of a carbon market can help to improve abatement and payoffs.

Table 2.4. Results for structure {M={USA, IND, ROW}, T={CHN, LAEX} }

		M without a carbon market		M with a carbon market	
		Mitigation in 2011 (Mton CO_2)	NPV of global payoffs over 100 years (billion $)	Mitigation in 2011 (Mton CO_2)	NPV of global payoffs over 100 years (billion $)
M	***USA***	761.5	7327.4	761.5	7515.2
(The	***IND***	1276.6	413.1	1276.6	463.2
coalition)	***ROW***	1187.6	1822.5	1187.6	1903.5
	Sum of M	3225.8	9562.9	3225.8	9881.9
The singletons	***CHN*** (trader)	894.8	750.5	1410.5	796.7
	LAEX (trader)	596.6	4037.7	467.6	4058.9
	EUR	783.6	12800.1	783.6	13242.0
	HAHI	289.8	11783.8	289.8	12180.8
	World	5790.6	38935.1	6177.3	40160.3

Second, with a carbon market T, some internally stable coalitions M, which are externally instable without a carbon market, are improved to be both internally and externally (i.e. fully) stable coalitions. These coalition structures are {EUR, IND, ROW} and {IND, HAHI, ROW}. In the presence of a carbon market, singletons face additional strategic options: to be carbon traders or outsiders of both M and T. We find that China prefers to be a trader over membership in a mitigation coalition. Tables 2.5a and 2.5b report the results of internally stable and fully stable M, which shows that due to the participation of CHN and LAEX in T, the CO_2 abatement of CHN increases from 894.8 Mton to 1410.5 Mton, because of the reduced emission levels by permits selling; In contrast, the abatement of LAEX decreases from 596.6 Mton CO_2 to 467.6 Mton CO_2 , because it imports permits instead of reducing CO_2 emissions. As a result, the absolute increase in abatement of CHN is larger than the absolute decrease in abatement of LAEX, the global level of abatement and payoffs of all global regions are improved. Without participation in T, the abatement levels of coalition $M = \{\text{IND}, \text{HAHI}, \text{ROW}\}$ (or $M = \{\text{EUR}, \text{IND}, \text{ROW}\}$) and singletons USA and EUR (or singletons USA and HAHI) remains the same as under the scenario without the carbon market, but their payoffs increase because of the positive externalities from the increased global abatement level.

Table 2.5a. Results for structure {M={IND, HAHI, ROW}, T={CHN, LAEX} }

		Internal stability of a single M		Full stability of M with a carbon market	
		Mitigation in 2011 (Mton CO_2)	NPV of global payoffs over 100 years (billion $)	Mitigation in 2011 (Mton CO_2)	NPV of global payoffs over 100 years (billion $)
M	***IND***	1416.9	416.6	1416.9	468.0
(The	***HAHI***	289.8	11765.7	289.8	12162.7
coalition)	***ROW***	1430.3	1841.6	1430.3	1924.7
	Sum of *M*	3137.0	14023.91	3137.0	14555.4
The singletons	***CHN*** (trader)	894.8	748.4	1410.5	794.5
	LAEX (trader)	596.6	4031.4	467.6	4052.7
	USA	601.5	9119.1	601.5	9436.2
	EUR	912.5	10396.4	912.5	10667.6
	Global	6348.5	38319.2	6529.1	39506.3

Table 2.5b. Results for structure {M: {EUR, IND, ROW}, T={CHN, LAEX}}

		Internal stability of a single M		Full stability of M with a carbon market	
		Mitigation in 2011 (Mton CO_2)	NPV of global payoffs over 100 years (billion $)	Mitigation in 2011 (Mton CO_2)	NPV of global payoffs over 100 years (billion $)
M	***EUR***	783.6	12647.1	783.6	13089.1
(The	***IND***	1372.3	414.3	1372.3	465.5
coalition)	***ROW***	1352.6	1838.9	1352.6	1921.8
	Sum of M	3508.5	14900.4	3508.5	15476.3
The singletons	***CHN*** (trader)	894.8	740.9	1410.5	787.0
	LAEX (trader)	596.6	3990.4	467.6	4011.6
	USA	601.5	9024.5	601.5	9341.7
	HAHI	337.0	9666.7	337.0	9917.0
	Global	5938.5	38322.8	6325.1	39533.5

Third, a carbon market has also negative impacts on stability. Without trade, the number of stable M is 28 totally, however, only 3 stable M are formed with trade. Lessmann et al. (2013) studied the carbon market mechanism in which the emission permits are traded between signatories and non-signatories of a climate agreement, showing that the extended market with non-signatories does not improve the incentive to join in the climate agreement compared with the case of permit trade only among signatories. Our findings about the negative effects of the carbon market are similar to Lessmann et al. (2013). One reason is that, with the existence of T, there are more potential strategy choices for each player $i \in N$, such as $(M_i = 1, T_i = 1)$, $(M_i = 1, T_i = 0)$, $(M_i = 0, T_i = 1)$ or $(M_i = 0, T_i = 0)$, which provides alternative options for each player to exploit the mitigation benefits but it also destabilises coalitions. Hence, some regions who chose to be members in a mitigation coalition in a setting without a carbon market will prefer to be singleton traders instead in a setting with a carbon market. China, as mentioned before, is a prominent example. Another drawback is reflected by the decrease of the indicator of success (the payoffs). Due to the establishment of an additional carbon market, the best-performing coalition under a single agreement system is destabilized, and the highest indicator of success can be reached by the stable coalition with a carbon market is only half of the best one obtained without a carbon market (see Table 2.A2). Although for a given coalition, the addition of carbon market has a positive influence on the indicator of success (see Tables 2.4, 2.5a and 2.5b).

In conclusion, for a given mitigation coalition, the presence of a carbon market could engage more regions with emissions mitigation, and thus the efficiency of global mitigation is enhanced with higher abatement and payoffs. However, at the same time, alternative strategic options offered by carbon trade destabilise the most efficient mitigation coalitions which are stable in the system with single mitigation agreement and without an open carbon market. It is also worthy to notice that the carbon market is not attractive for the mitigation coalition, which prefers to be a market outsider due to the free-riding incentives from the increased mitigation level.

2.6 Conclusion

We examined a four-stage game with sequential formation of a mitigation agreement and a carbon market with endogenous carbon endowments. The mitigation coalition M formed at stage 1 internalises the benefits of coalition members through the mitigation cooperation. The subsequently formed carbon market at stage 2 provides an opportunity of permits trading for all players if they join the market. In the next stage, the carbon market participants choose their permit endowments based on the endogenous permits trading model introduced by Helm

(2003). Permits trading among carbon traders and abatement levels are determined at the final stage.

We analysed this game using backward induction, and study the stability of the mitigation agreement under carbon trade and its efficiency by employing a refined version of the numerical model of STACO 3.0. Through comparison of different scenarios, we analysed the effects of a carbon market on the formation and efficiency of a mitigation agreement and find the following results.

First, in the presence of a carbon market with endogenous carbon permit endowments, players are faced with more strategic choices, which is beneficial to the singletons by choosing to join the market, but has an adverse influence on the formation of a mitigation coalition, since the members of some coalitions have an incentive to deviate from a mitigation coalition and become a carbon trader instead. This may destabilise the mitigation coalition. It implies that for the formation of ICAs, the alternative of joining a carbon market with endogenous carbon permit endowments can have a negative impact on coalition stability. Furthermore, from the results of stable coalition structures with a carbon market, this augmented strategy set destabilises the mitigation coalitions with higher indicator of success (see Table 2.A2) than the one obtained by including carbon trade, thus the efficiency and effectiveness of mitigation is reduced. Meanwhile, even with this dual-agreement system, the structure of coalition formation is still a partial participation with free-riding outsiders.

However, even though the negative influence of a carbon market's presence is obvious, it still could be beneficial to some mitigation coalitions by expanding the single coalition structure to improve the efficiency of some given suboptimal ICAs. For some specific coalitions which are stable with or without a carbon market, if a carbon market with endogenous carbon permit endowments is added, additional countries can be induced to increase mitigation efforts, which induces the Pareto improvement on abatement level and profitability of the given mitigation coalition (see Table 2.4). In Section 2.5, the comparison of the same stable mitigation coalitions with and without a carbon market shows that the total number of players involved in M and T is larger than the number of cooperative players in the absence of a carbon market, and the payoffs are also improved.

Third, our results also suggest some key regions in ICAs stabilization. Developing countries like India and China with low-cost mitigation options participate in both best-performing coalitions with or without carbon trade, which reinforces the importance of developing countries in an international agreement on emission reduction. Especially in the carbon market with endogenous carbon permit endowments, China bears a large quantity of abatement which contributes largely to global mitigation efficiency. This result reflects the need to engage developing countries with lower mitigation costs to cooperate with developed countries with higher abatement costs.

Finally, the analysis shows that a single mitigation coalition which prefers to be outsider of a carbon market, in contrast, cannot gain from opening its carbon market to outsiders, at least not under the condition of endogenous permit choice, because the negative transfers of buying emission permits and the benefits from free-riding surpass the gain from trading.

An important notice is that these conclusions are restricted to the investigation of strategic incentives into stability of ICAs without specific constraints on the choice of initial emissions permit. The formation and stability of ICAs could, however, be influenced by imposing constraints on permit endowments' choices. Therefore the relationship between stability of ICAs and putting specific caps on the issuing of emission permits is an interesting topic for future research.

2.7 Appendix

Appendix I

This Appendix shows the relationship between equilibrium price and the total number of initial permits choice at steady state. In a specific period t, the abatement costs for market participant $k \in T$ can be represented as Eq. (2.24) in Section 2.4, then by taking the derivative of Eq. (2.24) with respect to $q_{k,t}$, we obtain marginal abatement costs as

$$C'_{k,t} = \left(\alpha_k \cdot q_{k,t}^2 + \beta_k \cdot q_{k,t}\right) \cdot \chi_{k,t}, \qquad k \in T, \tag{2.A1}$$

From Eq. (2.4), we can get the specification of equilibrium price p_t^* as

$$p_t^* = C'_{k,t} = \left(\alpha_k \cdot q_{k,t}^2 + \beta_k \cdot q_{k,t}\right) \cdot \chi_{k,t}\ , \ \ k \in T, \tag{2.A2}$$

from which, we can get the function of after-trade abatement $q_{k,t}^*$ as

$$q_{k,t}^* = \frac{-\beta_k \cdot \chi_{k,t} + \sqrt{(\beta_k \cdot \chi_{k,t})^2 + 4\alpha_k \cdot \chi_{k,t} \cdot p_t^*}}{2\alpha_k \cdot \chi_{k,t}}, \ \ k \in T. \tag{2.A3}$$

Since the total number of initial permits at steady state is $\omega_t^* = \sum_{k\in T}\omega_{k,t}^* = \sum_{k\in T}\bar{e}_{k,t} - \sum_{k\in T} q_{k,t}^* = \bar{e}_t - \sum_{k\in T} q_{k,t}^*$, then we can get:

$$\omega_t^* = \bar{e}_t - \sum_{k\in T} \frac{-\beta_k \cdot \chi_{k,t} + \sqrt{(\beta_k \cdot \chi_{k,t})^2 + 4\alpha_k \cdot \chi_{k,t} \cdot p_t^*}}{2\alpha_k \cdot \chi_{k,t}}. \quad (2.A4)$$

This gives Eq. (2.26) in the main text.

Appendix II

Table 2.A1. Parameters in the 7-region STACO 3.0 model

	Share of global benefits	Parameter in abatement cost function	Parameter in abatement cost function	Uncontrolled Emissions (BAU) in 2011
Regions	s_j	α_j	β_j	Gton
USA	0.2263	0.00000189	0.02237106	7.0850
EUR	0.236	0.00000676	0.01950044	4.8917
CHN	0.062	0.00000243	-0.00076194	11.6240
IND	0.050	0.00001685	-0.00708971	3.2762
HAHI	0.2523	0.000071155	0.0411	7.5375
LAEX	0.1054	0.000018718	-0.00080322	4.7457
ROW	0.068	0.00000457	0.00984083	5.6603
World	$(\sum s_j = 1)$			

Table 2.A2. Stable coalitions M without a carbon market

Structure of stable M	NPV of global payoffs over 100 years (billion $)	Global mitigation in 2011in % of BAU emissions	Indicator of success
CHN, IND, HAHI, ROW	51892	18.7	72.6
EUR, CHN, IND, ROW	51501	19.5	71.4
USA, EUR, CHN	46874	18.5	56.8
USA, CHN, HAHI	46705	17.5	56.3
USA, CHN, ROW	45802	15.9	53.5
USA, CHN, LAEX	44841	15.7	50.5
USA, CHN, IND	44754	14.8	50.2
CHN, HAHI, LAEX	43937	15.9	47.6
EUR, CHN, LAEX	43859	16.6	47.4
CHN, LAEX, ROW	41314	13.4	39.4
CHN, IND, LAEX	40371	12.5	36.4
USA, EUR, ROW	39306	15.7	33.1
USA, IND, ROW	38935	12.9	31.9
USA, HAHI, ROW	38802	14.6	31.5
USA, EUR, IND	38618	14.2	30.9
USA, LAEX, ROW	38173	13.5	29.6
USA, IND, HAHI	38153	13.3	29.5
USA, IND, LAEX	37523	12.4	27.5
EUR, HAHI, ROW	37247	14.7	26.6
EUR, LAEX, ROW	37226	14.0	26.6
HAHI, LAEX, ROW	36919	13.4	25.6
USA, EUR, LAEX	36795	14.3	25.2
EUR, IND, HAHI	36686	13.3	24.9
EUR, IND, LAEX	36664	12.9	24.8
IND, HAHI, LAEX	36419	12.4	24.0
IND, LAEX, ROW	36139	11.4	23.2
USA, HAHI, LAEX	36116	13.2	23.1
EUR, HAHI, LAEX	34440	13.1	17.8

Chapter 3

International carbon trade with constrained allowance choices: Results from the STACO model*

International carbon markets are advocated in order to involve more countries in an agreement for the mitigation of greenhouse gas (GHG) emissions and to reduce the costs of mitigation. In this paper we develop a model where allowances are endogenously determined by each member of a carbon trade agreement, but with an exogenous constraint on the number of allowances per member. We use a global model to explore the incentives for regions to participate in such a carbon market and we examine its performance. To gain practical policy insights, we employ the STACO model, a numerically calibrated model with twelve world regions. Our results show that the stability and effectiveness of an international carbon market can be improved by imposing constraints on individual allowance choices compared to a carbon market without such constraints. Constraints on allowance choices reduce 'hot air' and increase global welfare and mitigation. When tightening the constraint *'broad but shallow'* agreements are replaced by *'narrow but deep'* ones. If the constraint is too tight, however, no stable carbon market exists.

* Published in Journal of Annals of Operations Research:
Yu S., Weikard H.-P., Zhu X., van Ierland E.C. (2016). International carbon trade with constrained endogenous allowance choices: Results from the STACO model, DOI: 10.1007/s10479-016-2126-3.

3.1 Introduction

Carbon emissions can be cost-efficiently reduced by means of carbon trade. The European Union Emission Trading Scheme (EU ETS) is an attempt to do this. Compared to a partial or regional carbon market, a global approach to carbon trade would engage all countries in emission mitigation. Moreover, inclusion of the major and low-cost emitters into the market could help to meet more ambitious mitigation targets and reduce abatement costs (Stern 2008; Behr et al. 2009; Flachsland et al. 2009). Yet, no global market for carbon has emerged so far. This can be explained by free-riding incentives to abstain from a global climate agreement (Barrett 1994). From the perspective of political efforts there are two approaches to a global emission trading system: the top-down approach based on government-to-government trading of emission allowances; and the bottom-up approach based on the linkage between regional emission trading systems (Stavins and Jaffe 2008; Behr et al. 2009; Flachsland et al. 2009). According to Flachsland et al. (2009), the top-down approach would generally cover a larger share of global emissions and is associated with larger mitigation efforts as compared to the bottom-up approach.

The possibility to meet an emission reduction target by means of a government-to-government emission trading system was firstly established and specified by the Marrakesh Accords in 2001 based on the Kyoto Protocol. However, this trading system only includes the developed countries (listed in Annex I of the Kyoto Protocol). Given that growing shares of global emissions stem from emerging (non-Annex I) economies, like China and India, the effectiveness of a partial international trading system supported by the Marrakesh Accords and the Kyoto Protocol can be enhanced. As a project-based trading system, the Clean Development Mechanism (CDM) allows the Annex I countries to get tradable emission reduction credits when investing in emission-reduction projects in developing countries (UNFCCC 1998). A limitation of the CDM is that only the Annex I countries are committed to the Kyoto mitigation targets, whereas developing countries are not committed to any mitigation.

In a market for carbon emission allowances countries with relatively high abatement costs would be buyers and have incentives to join a carbon market as the market offers cheaper abatement options. Countries with relatively low abatement costs would be sellers and could gain from earning revenues. In a carbon market with unconstrained endogenous allowance choices, as studied by Helm (2003), the motivation to raise revenues from carbon trading results in excessive allowance choices, so-called 'hot-air'. Such a carbon market is then characterised by modest emission reductions. Moreover, the stability of a carbon market with open membership could be undermined by incentives for participation that stem from selling carbon emission allowances. Therefore, imposing a constraint on allowance choices might not only mitigate the hot-air effect but can also help to stabilise a carbon market by avoiding excessive participation of potential sellers. However, a constraint on allowance choices can also generate free-riding incentives because the improved global abatement resulting from limiting carbon emission allowances will increase the payoffs of non-signatories. It is therefore important to study how exogenous constraints on allowance choices change the incentives to participate in an international carbon market and its environmental effectiveness. We address this problem by modelling a top-down approach to an international carbon market where emission allowances are traded between governments.

Stevens and Rose (2002) studied a restricted carbon market with a constraint on the volume of carbon transactions, i.e. purchases and sales of emission allowances, and showed that abatement costs would be increased due to the carbon trade restrictions. Rehdanz and Tol (2005) analysed the impacts of regulation imposed on a bilateral carbon market. They assume that the carbon buying country will suffer higher damages from GHG emissions. Therefore, the carbon buying country can strategically and unilaterally set stricter abatement targets for its own emissions aiming to reduce its carbon allowance imports and limiting emission permits issued by the selling country. Their research shows that the regulation of the quantity of emission permits makes both countries worse off if the regulation adopted by the buying country is strict, as this reduces cost-savings from trade. Altamirano-Cabrera and Finus (2006) consider uniform emission reductions to define tradable quota. They study the impact of restricted carbon trade on the formation and efficiency of climate coalitions, but they do not consider allowance choices. Carbone et al. (2009) apply Helm's idea of a carbon market with endogenous allowance choices in a calibrated general equilibrium model. They analyse participation incentives and environmental effectiveness of international carbon trade agreements. In order to mitigate the hot-air effect Carbone et al. (2009) consider a setting where incumbent members of the carbon market may block entry of additional potential market participants, i.e. they only consider internal stability of the carbon market. In this paper we explore another option. While we maintain the idea of a carbon market with endogenous allowance choice, we use an open-membership model, in line with most currently existing international environmental agreements. In our model the hot-air effect is mitigated by constraining the allowance choice. The constraint is exogenous to our model. Hence, we consider it to be part of the agreement that is “on the table” ready to be signed.

We explore the incentives to join a carbon market in a two-stage non-cooperative game. In the first stage, regions decide simultaneously on their market participation. In the second stage signatories can choose their allowances and then trade. However the carbon trade agreement obliges all signatories to accept a constraint on their allowance choice that we model as a fraction of the business-as-usual (BAU) emissions. This setting is in line with the ‘cap-and-trade’ system of the EU ETS where caps were set relative to historical emission levels. BAU emission levels reflect historical emission levels, i.e. carbon emissions before any (unilateral) climate policies were adopted. Historical emission levels have also played a role as reference points in climate negotiations. Another reason for using BAU emissions levels as our base line is that, unlike Nash equilibrium levels, the BAU levels are exogenous to our model. Since it is still interesting to explore the setting when allowance constraints are tied to Nash-emissions levels, we provide results for this case in a sensitivity analysis. Generally we assume that non-signatories (or singletons) cannot participate in international carbon trade and, thus, they adopt their own carbon abatement policies. We examine cartel stability, i.e. we assume a single international carbon market which is stable if no signatory has an incentive to leave the market and no singleton has an incentive to join.

Intuitively, a carbon market with a constraint on allowance choices can be more effective in terms of emission reductions compared to an unconstrained market. We show this by employing the STACO model. STACO specifies business-as-usual emission paths and emission abatement costs and benefits functions for twelve heterogeneous world regions (Nagashima et al. 2011; Dellink et al. 2015). We use the model to identify stable carbon trade agreements. In particular we show that by tightening the constraint on allowance choices, the global mitigation level in a stable carbon market can be improved. Regarding the welfare effects, there are two main consequences. Firstly, by imposing a constraint on allowance choices, the benefits from global abatement can be increased, especially for the countries with relatively high marginal benefits from global abatement. Secondly, obviously, constraints on allowance choices reduce the supply of emission allowances in the carbon market. This will reduce hot air, drive up the carbon price and thereby reduce the benefits for carbon buyers. The revenues of carbon sellers could also be reduced as only a limited number of emission allowances can be sold. Furthermore, in our setting with asymmetric regions, individual welfare effects will differ per region due to the differences in marginal abatement costs and benefits.

The main idea of our paper is to investigate the impact of allowance choice constraints on participation incentives in a carbon market. The impact of the constraint is analysed by varying the level of the constraint parameter in our simulation analysis. We do not make a claim about which level of the constraint parameter would be chosen in the pre-negotiations to a trade agreement, i.e. we are not endogenising the constraint. In our approach the optimal level (in terms of global welfare) of the constraint is identified from numerical results from the STACO model.

The paper is structured as follows. Section 3.2 describes our two-stage game of the formation of an international carbon market. In Section 3.3, we introduce the model with and without allowance choice constraints and we analyse the two-stage game by backward induction. The numerical analysis is implemented, and the results are presented and discussed in Section 3.4. The final section summarizes the main findings and discusses policy implications.

3.2 Formation of an international emission trade agreement

A standard two-stage coalition formation game is applied to study an international carbon market with heterogeneous regions. Each region is characterised by its abatement cost and benefit functions. The set of all regions is denoted by N. An individual region is indexed by j with $j = 1, \dots, n$.

At stage 1, a membership game is played. All regions $j \in N$ simultaneously and non-cooperatively choose whether or not to join a proposed carbon market. The choice set is defined as $\sigma_j = \{0,1\}$. If $\sigma_j = 1$, then j joins the market. If $\sigma_j = 0$, then j does not join and remains a singleton. Countries decide upon their membership by anticipating the welfare impacts of the allowance choices and the ultimate abatement level. We refer to the set of regions who join the market as traders T.

At stage 2, with a given set of traders T, every trader $j \in T$ chooses initial allowances, denoted by ω_j, subject to a constraint and chooses abatement, denoted by q_j, depending on carbon trade. The constraint specifies the maximum level of allowances that each market participant can choose, denoted by ω_j^{max}. Specifically, the individual maximum allowances

ω_j^{max} are a fraction of the BAU emissions $\bar{e}_j$ ($j \in T$) such that $\omega_j^{max} = \alpha\bar{e}_j$. In our model the parameter $\alpha \in [0, 1]$ is the same for all carbon traders, but the maximum allowance choices ω_j^{max} ($j \in T$) are different across traders since their BAU emissions $\bar{e}_j$ differ. As the mitigation target becomes stricter, the value of α decreases. The strictest possible constraint, $\alpha = 0$, refers to a carbon-free economy. However, in our numerical simulations we do not consider constraints on allowance choices that are stricter than what the social optimum requires. In fact, our results from the STACO model in Section 3.4 show that no stable carbon market can be found when the constraint parameter is lower than 0.74 (see Table 3.3 in Section 3.4.2).

Following Helm's (2003) endogenous allowance choice model, each trade participant determines its best response allowance choice and after-trade abatement by solving the following problem:

$$\max_{\omega_j, q_j} \quad \pi_j = B_j\left(\sum_{j \in T} \bar{e}_j - \omega + \sum_{j \in N-T} q_j\right) - C_j(q_j) + p(\omega) \cdot \left(\omega_j - (\bar{e}_j - q_j)\right), \; j \in T, \quad (3.1)$$

subject to

$$0 \leq \omega_j \leq \alpha\bar{e}_j, \qquad with\, \alpha \in [0, 1], q_j \geq 0.$$

The global abatement level is $q = \sum_{j \in T} \bar{e}_j - \omega + \sum_{j \in N-T} q_j$. In the objective function (Eq. (3.1)), the total allowances are $\omega \equiv \sum_{j \in T} \omega_j$, which also represents the total emissions in the carbon market T. The carbon price is a function of emission allowances ω in the carbon market, denoted by $p(\omega)$ with $p'(\omega) \leq 0$. Carbon price $p(\omega)$ is uniform for all trade participants. In line with the functional forms in the STACO model, we assume abatement benefits $B_j(q)$ that are linear in the total abatement with $B_j'(q) > 0$ and $B_j''(q) = 0$. Abatement costs $C_j(q_j)$ are strictly convex in individual abatement with $C_j'(q_j) > 0$ and $C_j''(q_j) > 0$. Emissions are denoted by $e_j = \bar{e}_j - q_j$ corresponding to the abatement level q_j. Note that the functional forms of $B_j(q)$ and $C_j(q_j)$ result in a concave net benefit function $B_j(q) - C_j(q_j)$, which assures that the optimal solution to the problem (1) is uniquely determined.

At this second stage, the non-traders $j \notin T$ choose their optimal mitigation level by solving

$$\max_{q_j} \quad \pi_j(q_j) = B_j(q) - C_j(q_j), \ \ j \notin T. \tag{3.2}$$

Because of the linear form of the abatement benefits function, singletons have a dominant strategy implying that the abatement level of any singleton is not influenced by the carbon market.

3.3 Model and the theoretical analysis

In this section we describe the sub-game perfect Nash equilibria (SPNE) of the game using backward induction.

Stage 2: Equilibrium choices of allowances and abatement

At the second stage, given the set of trade participants T and a constraint $\omega_j \leq \alpha\bar{e}_j$ on individual allowance choices, the initial emission allowances and the after-trade abatement levels are chosen. The maximisation problem (3.1) gives the following Lagrangian function:

$$L(\omega_j; q_j; \gamma_j) = \pi_j + \gamma_j(\alpha\bar{e}_j - \omega_j), j \in T. \tag{3.3}$$

In (3.3) γ_j is the Lagrangian multiplier for individual emission allowance choices. By taking the derivatives of Eq. (3.3), the first order conditions for ω_j, q_j and γ_j are derived as follows:

$$\frac{\partial L}{\partial \omega_j} = \frac{\partial \pi_j}{\partial \omega_j} - \gamma_j = p'(\omega^*)(\omega_j^* - e_j^*) + p - B_j' - \gamma_j = 0, \tag{3.4}$$

$$\frac{\partial L}{\partial q_j} = -C_j' + p = 0, \tag{3.5}$$

$$\frac{\partial L}{\partial \gamma_j} = \alpha\bar{e}_j - \omega_j \geq 0, \tag{3.6}$$

$$\gamma_j(\alpha\bar{e}_j - \omega_j) = 0, \gamma_j \geq 0. \tag{3.7}$$

From Eq. (3.5) we conclude that the equilibrium carbon price equals to the marginal abatement cost $p(\omega^*) = C_j'(q_j^*)$. Rewriting Eq. (3.4) leads to the following equilibrium condition:

$$B_j' - C_j' = p'(\omega^*)\left(\omega_j^* - (\bar{e}_j - q_j^*)\right) - \gamma_j. \tag{3.8}$$

When the allowance choice constraint $\alpha\bar{e}_j$ is not binding, then the equilibrium conditions will be identical to the equilibrium condition for unconstrained carbon markets where the shadow value $\gamma_j = 0$ in Eq. (3.8). This shows that with unconstrained allowance choices, the marginal revenues (either negative or positive) from carbon trade $p'(\omega^*)\big(\omega_j^* - (\bar{e}_j - q_j^*)\big)$ are equal to the net marginal abatement benefits $B_j'(q^*) - C_j'\big(q_j^*\big)$. Compared to the unconstrained carbon markets, Eq. (3.8) shows that with constraints, each carbon trader's marginal abatement net benefits do not only depend on the marginal trade revenues $p'(\omega^*)\big(\omega_j^* - (\bar{e}_j - q_j^*)\big)$, but are also affected by the shadow value of the allowance choices if the constraint is binding. With a constraint on allowance choices, the shadow value γ_j represents the unavailable marginal gains due to the constraint on j's allowances. Substituting $C_j'(q_j^*)$ with $p(\omega^*)$ in Eq. (3.8) we obtain the carbon price

$$p(\omega^*) = B_j' - p'(\omega^*)\left(\omega_j^* - \left(\bar{e}_j - q_j^*\right)\right) + \gamma_j. \tag{3.9}$$

Eq. (3.9) implies that the carbon price will be impacted by the constraint. The more stringent the constraint is, the higher is the shadow price of allowances and the more valuable are emission allowances.

Now consider that the value of the constraint parameter α decreases such that the constraint is binding for all carbon traders and, hence, the optimal choice of initial allowances is $\omega_j^* = \alpha\bar{e}_j$. Then the size of the market is $\omega^* = \alpha\sum_{j\in T}\bar{e}_j$. Any further decrease of α further reduces the optimal allowance choices ω_j^* and, hence, ω^*. Reduced allowance choices require increased after-trade abatements $q_j^* = q_j^*(p(\omega^*))$, higher marginal abatement cost and, since $p(\omega^*) = C_j'(q_j^*)$, also the carbon price $p(\omega^*)$ is higher.

It is interesting to consider the implications of the Kuhn-Tucker condition $\frac{\partial\pi_j}{\partial\omega_j} - \gamma_j = 0$; see Eq. (3.4), rewritten in (3.8). When the shadow value of allowances $\gamma_j = 0$, the constraint is not binding and carbon traders j ($j \in T$) will choose their optimal allowances as $\omega_j^* < \alpha\bar{e}_j$. Then the equilibrium condition will be identical to the equilibrium in an unconstrained carbon market. This can happen when the constraint is lax, especially to the carbon buyer regions with high marginal abatement benefits and costs. The numerical results shown in Table 3.3 confirm that carbon buyer regions like USA and Japan choose non-binding levels of allowances with a lenient constraint. However, when the constraint is strict enough, it will be binding such that the shadow value $\gamma_j > 0$ in Eq. (3.8). If the allowance choice constraint is binding, traders j ($j \in T$) choose their optimal allowances as $\omega_j^* = \alpha \bar{e}_j$. Since $\frac{\partial \pi_j}{\partial \omega_j} = \gamma_j > 0$ it can be concluded that the gains of all carbon traders decrease with the tightening of the constraint on allowance choices. Consequently, the participation incentives decrease and thus there is no carbon market when the constraint is too strict. In this case the internal stability condition, specified below, is violated.

Finally, to conclude the analysis of stage 2, we still need to consider the behaviour of the singletons. As they cannot participate in the market, they decide their mitigation levels by maximising the payoffs specified in Eq. (3.2). Their best response is characterised by $C_j'(q_j^*) = B_j'(q^*)$.

Stage 1: Membership choice

At stage 1, all players make their membership decisions considering how the stage-2 game will be played. The Nash equilibria of the membership game correspond to cartel stability (d'Aspremont et al., 1983; Barrett 1994). Hence, in an equilibrium carbon market satisfying internal and external stability, no trade participant has an incentive to leave and no singleton has an incentive to participate; see conditions (3.10) and (3.11) below. For a carbon market T with constraint parameter α, we introduce a partition function $V_j(T;\ \alpha)$ to represent the payoffs of trade participants j ($j \in T$) and singletons j ($j \notin T$) as a function of the set of traders. An equilibrium carbon market is formed if the following internal and external stability conditions hold:

Internal stability: $V_j(T;\ \alpha) \geq V_j(T\backslash\{j\};\ \alpha), j \in T,$ (3.10)

External stability: $V_j(T;\ \alpha) \geq V_j(T \cup \{j\};\ \alpha), j \notin T.$ (3.11)

3.4 Simulation analysis and results

To illustrate the consequences of imposing constraints on individual allowance choices for the stability and the performance of an international carbon market, we conduct a simulation analysis based on our two-stage game employing the STACO model. A detailed description of the numerical approach is presented in section 3.4.1. To compare and analyse the differences in terms of stability and performance of international carbon markets with and without allowance choice constraints, we firstly examine the base scenario of an unconstrained carbon market. Results are provided in section 3.4.2. The results for constrained carbon markets are presented and discussed in section 3.4.3. In section 3.4.4., as a sensitivity analysis, we provide results for a scenario where the non-cooperative Nash equilibrium emissions are used as a baseline for the allowances constraint.

3.4.1 Simulations employing the STACO model

Our simulation is performed by employing the STACO model, which is an integrated assessment model connecting GHG emissions with abatement costs and economic evaluations of climate damages for twelve different world regions: United States (USA), Japan (JPN), European Union-27 & EFTA (EUR), Other High Income countries (OHI), Rest of Europe (ROE), Russia (RUS), High Income Asian countries (HIA), China (CHN), India (IND), the Middle East countries (MES), Brazil (BRA) and Rest of the World (ROW). In STACO, the economic evaluation with respect to the payoff assessment is specified by the comparison between abatement costs and benefits. Focusing on the establishment of an international carbon market, we modify carbon traders' payoff functions by including the carbon trade effects into the original payoff function in STACO under different constraints on allowance choices, as shown in Eq. (3.1). Considering inertia and the long term effects in the climate system, the STACO model evaluates GHG emission mitigation with projected baseline emissions $\bar{e}_j$ for a horizon of 100 years. Given the participation choice in the initial period, each region determines optimal abatement levels q_j^* strategically in every time period. For a full specification of the latest version of the STACO model (STACO 3.0) the reader is referred to Nagashima et al. (2011) and Dellink et al. (2015).

We use the numerical computing software MATLAB to do the calculations and to derive the numerical solutions. The approach to solving the numerical model is implemented in steps: Firstly, we assign a number to all possible non-trivial carbon markets ($|T| \geq 2$). With 12 regions, the total number of non-trivial carbon markets is $2^{12} - 12$. In the second step, we calculate the equilibrium allowance choices, after-trade abatement and payoffs of each trader and for every possible market. We also calculate abatement levels of the singletons. Then we repeat this step for a tighter constraint on allowance choices, i.e. we lower the value of α stepwise (with step size $\Delta\alpha = 0.02$) from 1 to the percentage of the social optimum emissions. From these results, we can observe the changes in abatement and payoffs of all players associated with different constraints. Based on the results from the second step, the equilibrium carbon markets can be identified and internal and external stability can be checked. Finally, from the performances and stability of the equilibrium carbon markets, we can identify an optimal exogenous constraint.

3.4.2 Results for carbon markets with unconstrained allowance choices

In an unconstrained carbon market the initial emission allowances can be chosen arbitrarily by trade participants. We first look at the stability of all possible non-trivial carbon markets. Our result shows that no equilibrium carbon market is simultaneously satisfying internal and external stability conditions under unconstrained allowance choices. However, if we only consider internal stability, then there are 36 internally stable carbon markets. The internally stable carbon markets are listed in Table 3.1, which reports the 12 best-performing carbon markets in terms of the global NPV (net present value) over 100 years. From the last column of Table 3.1, we can see that with arbitrary allowance choices, carbon market participation is attractive to all potential carbon sellers which are characterized by low marginal abatement benefits and low marginal abatement costs. Full stability is undermined by the incentives for singletons to join, i.e. external stability is violated. Specifically, regions with low marginal benefits (i.e. see Table 3.A1 in Appendix) from global abatement are highly motivated to join, for example OHI, ROE, RUS, HIA, IND, MES, BRA and ROW. The underlying reason is that regions with lower marginal benefits from global abatement have less incentive to reduce their emissions, hence their main motivation to join a carbon market is seeking revenues from carbon sales. This finding indicates that it will be difficult to establish an unconstrained carbon market under open membership. A constraint on the allowance choices could limit the participation incentives that are arising from revenue seeking behaviour which, in turn, results in hot air. A constraint on allowance choices may improve external stability.

Table 3.1 shows that global mitigation and welfare of the internally stable carbon markets are higher compared to the non-cooperative Nash equilibrium, but are much lower than the levels in the social optimum (see the first and second row of Table 3.1). This can be attributed to non-cooperative strategies of abatement choice and limited participation. As we can see from the second to the last row of Table 3.1 the global payoffs and abatement levels achieved under carbon trade are significantly higher than the no-trade Nash equilibrium. The welfare enhancement through carbon trade is mainly due to the cost-savings from cheaper abatement options. Exploiting cheaper abatement options increases global mitigation. Regions with high marginal abatement costs (MAC) (see Figure 3.A1 in Appendix) like JPN and EUR are motivated to abate more than their Nash equilibrium levels because of their higher shares of global abatement benefits. At the same time, a region like CHN which sells carbon without producing hot-air also contributes to the global abatement compared to the no-trade outcome. Hence, with carbon trade, the global emission reductions are larger and cheaper than in the no-trade Nash equilibrium. Both traders and non-traders can benefit from the positive externality of the increased global mitigation.

The best-performing market is formed by two countries with different properties: EUR, which is characterized by high MAC, and CHN, which has the lowest MAC of all regions. In the best performing market, the MACs of the two traders are equalised at a carbon price of 4.74 $/tonCO_2$. This price is much lower than the MAC of EUR in the non-cooperative equilibrium and thus EUR benefits a lot from emission trading. In Carbone et al. (2009), the best market is also composed of EUR and CHN, which is comparable to our result. It is worth noticing the different carbon prices resulting in these different market structures as shown in column 4. Prices in the three best performing markets are higher than in other carbon markets. This is due to the number of participants with low MAC in the markets with a better performance. It is obvious that the carbon markets with lower carbon prices like {JPN, CHN, IND} and {EUR, CHN, IND} are formed by more regions with low MAC compared to the carbon markets like {EUR, CHN} and {JPN, CHN}.

As shown by Helm (2003) and Carbone et al. (2009), the arbitrary choice of emission allowances can cause hot air in unconstrained carbon markets. Our results also confirm the existence of the hot-air effect. Table 3.2 reports the hot air effects that we find in eight of the twelve best-performing carbon markets listed in Table 3.1. Only the four best performing carbon markets do not show hot air effects. As explained in Carbone et al. (2009), hot air is not an issue for China as it prefers to maintain a relatively higher carbon price by reducing allowances. Hence the markets where China is the only carbon seller do not show hot air effects. For the carbon market {JPN, CHN, IND} the allowances supply by the world's two biggest carbon sellers China and India can easily satisfy the small demand of Japan, without supplying an amount beyond their BAU emissions. So, also in this case hot air is avoided. Comparing hot-air and global abatements, see columns 1 and 2 in Table 3.2, we find that the abatement is inversely associated with the magnitude of hot-air. This observation carries over to global payoffs which are also inversely related to the magnitude of hot-air. These results underline the negative impact of hot-air on abatement and welfare. According to Helm (2003), it is possible that the total allowances chosen by all carbon traders exceed their non-cooperative emissions level under certain conditions. However, in our simulations total allowances of all carbon traders are lower than their non-cooperative emissions. This can be seen by comparing the no-trade Nash emissions level and the allowance choices in a carbon market in Table 3.2. The reason is that although some traders choose allowances above the BAU level, other traders (i.e. EUR and JPN) who have large marginal benefits of global abatement will choose lower allowances than their no-trade Nash emissions level to offset the negative influence of others' excessive allowance choices.

Table 3.2 shows that the hot air effects is largest in the carbon market {EUR, CHN, BRA}, caused by excessive allowance choice of BRA. The driving force is Brazil's high marginal abatement costs, so it has an incentive to drive down the carbon price by increasing the amount of emission allowances; at the same time, Brazil's marginal benefits are quite low (see Table 3.A1 in Appendix) and, hence, its incentives to reduce emissions are limited. As shown in the third row in Table 3.2, our results confirm that ROE, consisting of mainly the former Soviet Union countries, is one of the largest sources of hot air. In addition to its low benefits share from global abatement, the slow economic growth also reduces ROE's demand for GHG emission allowances. The reason why MES produces hot air can be understood because it is the largest exporter of fossil fuels. MES seeks to decrease the carbon price and thus to increase the carbon demand which has a positive influence on fossil fuel exports. The hot air effects generated in other three carbon markets shown in Table 3.2 are caused by regions like HIA, OHI and IND respectively. The common reason for hot air created in these three markets is their relatively low benefits from global abatement (see Table 3.A1 in Appendix) and high revenues from carbon sales.

In summary, our numerical analysis confirms the advantage of carbon markets in terms of global welfare and mitigation as compared to the no-trade Nash equilibrium. However, without a constraint on allowance choices it is difficult for a carbon market with open membership to satisfy external stability, because it is easily destabilized by market entrants who seek to raise revenues from carbon sale. The lower a region's marginal damages of GHG emissions, the more allowances will be chosen. Excessive allowance choices result in an inefficient carbon market.

Table 3.1. Results for the top twelve best-performing internally stable carbon markets with unconstrained allowance choices

Market structures with internal stability	NPV of Global payoffs over 100 years (billion $)	Global mitigation in 2011 in % of BAU emissions	Price of emission allowance in 2011 ($/ton CO_2)	Singleton regions with incentives to join the carbon markets
No-trade Nash equilibrium (All singletons)	8446.39	5.38	------	All
The social optimum equilibrium	29559.56	28.75	30.52[a]	------
{EUR, CHN}	9682.73	7.00	4.74	OHI, ROE, RUS, HIA, IND, MES, BRA, ROW
{JPN, CHN}	9556.75	7.07	3.95	OHI, ROE, RUS, HIA, IND, MES, BRA, ROW
{USA, CHN}	9551.28	6.69	3.53	OHI, ROE, RUS, HIA, IND, MES, BRA, ROW
{JPN, CHN, IND}	9416.88	7.03	2.75	OHI, ROE, RUS, HIA, MES, BRA, ROW
{EUR, CHN, IND}	9399.96	6.86	3.28	OHI, ROE, RUS, HIA, MES, BRA, ROW
{JPN, HIA, CHN}	9143.16	6.74	2.91	OHI, ROE, RUS, IND, MES, BRA, ROW
{JPN, OHI, CHN}	9122.71	6.73	2.90	ROE, RUS, HIA, IND, MES, BRA, ROW
{JPN, ROE, CHN}	9113.74	6.84	2.74	OHI, RUS, HIA, IND, MES, BRA, ROW
{JPN, CHN, MES}	9087.21	6.70	2.84	OHI, ROE, RUS, HIA, IND, BRA, ROW
{EUR, ROE, CHN}	9060.98	6.63	3.26	OHI, RUS, HIA, IND, MES, BRA, ROW
{JPN, CHN, BRA}	9052.53	6.68	2.78	OHI, ROE, RUS, HIA, IND, BRA, ROW
{EUR, CHN, BRA}	8999.96	6.45	3.30	OHI, ROE, RUS, HIA, IND, BRA, ROW

a: Equalised marginal abatement costs in the social optimal equilibrium.

Table 3.2. Hot-air effects in carbon markets with unconstrained allowance choices

Structure of internally stable markets	Magnitude of hot-air in 2011 ($MtonCO_2$)	Global mitigation in 2011 in % of BAU emissions	Choices of emission allowance in 2011 ($MtonCO_2$)											
			USA	JPN	EUR	OHI	ROE	RUS	HIA	CHN	IND	MES	BRA	ROW
BAU emissions	---	0	7085.96	1386.55	4891.74	1411.23	1503.11	2234.26	2591.49	11623.98	3276.2	991.67	2181.62	5660.27
No-trade Nash emissions	---	5.38	6801.33	1329.58	4488.58	1393.43	1431.75	1967.92	2575.2	10952.99	2811.19	982.39	2173.14	5517.41
The social optimum emissions	---	28.75	5845.02	1163.70	3761.62	1021.15	888.45	1329.79	2059.67	7883.93	1690.18	521.02	1851.04	3931.46
{EUR, CHN, BRA}	628.33	6.45	0	0	3391.85	0	0	0	0	10930.92	0	0	**2809.95**	0
{EUR, ROE, CHN}	544.89	6.63	0	0	3294.35	0	**2048.00**	0	0	10971.09	0	0	0	0
{JPN, CHN, BRA}	458.64	6.68	0	378.87	0	0	0	0	0	10853.03	0	0	**2640.26**	0
{JPN, CHN, MES}	434.31	6.70	0	401.92	0	0	0	0	0	10848.83	0	**1425.98**	0	0
{JPN, HIA, CHN}	393.73	6.74	0	409.12	0	0	0	0	**2985.22**	10854.52	0	0	0	0
{JPN, OHI, CHN}	390.97	6.73	0	412.86	0	**1802.20**	0	0	0	10852.22	0	0	0	0
{JPN, ROE, CHN}	379.39	6.84	0	285.90	0	0	**1882.50**	0	0	10891.83	0	0	0	0
{EUR, CHN, IND}	134.66	6.86	0	0	3122.58	0	0	0	0	11055.56	**3410.86**	0	0	0

Note: Numbers in bold face indicate hot air.

3.4.3 Results for carbon markets with constraints on allowance choices

In this part, we focus on the numerical analysis of a carbon market with a constraint on the allowance choices. As described in Section 3.3, allowance choices are limited to a fraction α of each trade participant's BAU emissions. Table 3.3 reports the stable carbon markets that we find for different levels of α. Compared to the unconstrained scenario, stability of and participation in the constrained carbon market are improved; with constraints we find 16 stable carbon markets and the maximum number of trade participants is increased to five (see last column of Table 3.3). As shown in Table 3.1, under unconstrained allowance choices, the potential carbon sellers have incentives to join (i.e. ROE, RUS, IND, BRA and ROW). With constraints on allowance choices, these singleton regions are discouraged to join the carbon trade and the degree of discouragement increases with the strictness of the constraints. From columns 1 and 4 of Table 3.3 we can see that when α falls from 0.98 to 0.88, regions like ROE, RUS and IND are still willing to join; but when the value of α is decreased further, only RUS is motivated to join. Because ROE and IND are the main source of hot-air (as shown in Table 3.2), their marginal benefits from global abatement are much lower compared to Russia's (see marginal benefits in Table 3.A1). When the constraint on allowance choices is too strict to make profits from selling carbon, ROE and IND prefer not to join the market. Table 3.3 shows that with decreasing of α from 0.84 to 0.74, only the carbon market {JPN, RUS} still remains to be stable. This is because both JPN and RUS have incentives to reduce emissions through carbon trade due to relatively higher marginal abatement benefits, especially JPN. At the same time, even under a stricter constraint, JPN can still find cheaper abatement options from carbon trade, while RUS also can earn revenues from selling carbon. However, no stable carbon market can be found when α is lower than 0.74. Note that the social optimum requires that global emission would fall to 71% of BAU emissions; see Table 3.1.

Table 3.3 displays that a stable carbon market will be the largest (in terms of the number of traders) when the constraint is modest, e.g. $\alpha = 0.9$ and 0.88. Moreover, multiple equilibrium carbon markets can emerge under modest constraints, e.g. $\alpha = 0.94$, 0.92 and 0.88. However, when the constraint becomes stricter, e.g. $\alpha < 0.88$, the number of stable markets and their size decrease. The list of the equilibrium carbon markets under different constraints, shown in the last column of Table 3.3, also shows that when $\alpha < 0.88$, the stability of the equilibrium carbon market {JPN, RUS} is robust to varying the allowance constraint. This robustness implies that for regions like JPN and RUS, the ranges of cost savings in abatement and earnings from carbon sale are so large that they can still benefit from carbon trade under more stringent constraints. It is interesting to observe the relationship between the actual allowance choices and the constraint. As shown in column 2 of Table 3.3, under less strict constraints, e.g. from $\alpha = 0.98$ to 0.88, the actual allowance choices of carbon buyers (i.e. USA and JPN) with higher marginal abatement benefits, are generally non-binding. Carbon sellers (i.e. IND, ROE and RUS) who have lower marginal benefits from global abatement but can raise revenues from carbon sale, prefer to choose allowances at the binding levels (shown by bold numbers in Table 3.3). However, when the constraint is becoming stricter, e.g. $\alpha < 0.88$, it becomes binding for all traders.

Table 3.3 also shows that, a global market involving all world regions is difficult to realise. However, at least some larger GHG emitters like USA, IND and RUS can be included in a carbon trade agreement. The non-existence of a stable global carbon market indicates that the constraints imposed on the allowance choices can stabilise the carbon market to a certain degree but cannot completely overcome free-riding incentives. It is worth noting that CHN does not join any carbon market satisfying internal-external stability under different constraints, in contrast to the observation that CHN is part of the internally stable unconstrained markets discussed in the previous subsection. To see why this is the case notice that the markets listed in Table 3.1 are externally unstable because others would like to join, create hot air, and thus destabilise the (enlarged) market. Although hot air is ruled out in the constraint market, a similar mechanism is at work. Consider a lax constraint, e.g. $\alpha = 0.98$. Here CHN would not join $\{$ROE, HIA, MES$\}$ because ROE would then increase allowances. For stricter constraints the reason for the absence of China is that market entry requires a tough restriction of allowances such that potential revenues from sales of allowances are overcompensated by higher abatement costs.

Table 3.3. Stable carbon markets under different constraints

Constraint parameter α	Allowance choices in 2011 in % of BAU emissions [b]								Carbon markets with full stability
	USA	JPN	ROE	RUS	HIA	IND	MES	ROW	
0.98	-----	-----	0.97	-----	**0.98**	-----	**0.98**	-----	{ROE, HIA, MES}
0.94	-----	0.93	**0.94**	-----	-----	-----	-----	-----	{JPN, ROE}
0.94	0.92	0.91	**0.94**	-----	-----	**0.94**	-----	-----	{USA, JPN, ROE, IND}
0.94	-----	0.92	**0.94**	-----	-----	0.93	-----	**0.94**	{JPN, ROE, IND, ROW}
0.92	-----	0.91	**0.92**	-----	-----	-----	-----	-----	{JPN, ROE}
0.92	0.91	0.86	**0.92**	-----	-----	0.91	-----	-----	{USA, JPN, ROE, IND}
0.9	**0.90**	0.88	**0.90**	**0.90**	-----	**0.90**	-----	-----	{USA, JPN, ROE, RUS, IND}
0.88	-----	0.81	-----	-----	-----	**0.88**	-----	-----	{JPN, IND}
0.88	**0.88**	-----	-----	**0.88**	-----	**0.88**	-----	-----	{USA, RUS, IND}
0.88	**0.88**	0.86	**0.88**	**0.88**	-----	**0.88**	-----	-----	{USA, JPN, ROE, RUS, IND}
0.84	-----	**0.84**	-----	**0.84**	-----	-----	-----	-----	{JPN, RUS}
0.82	-----	**0.82**	-----	**0.82**	-----	-----	-----	-----	{JPN, RUS}
0.8	-----	**0.80**	-----	**0.80**	-----	-----	-----	-----	{JPN, RUS}
0.78	-----	**0.78**	-----	**0.78**	-----	-----	-----	-----	{JPN, RUS}
0.76	-----	**0.76**	-----	**0.76**	-----	-----	-----	-----	{JPN, RUS}
0.74	-----	**0.74**	-----	**0.74**	-----	-----	-----	-----	{JPN, RUS}
0.72	-----	-----	-----	-----	-----	-----	-----	-----	None

b: The ratio is calculated as $\omega_j^*/\bar{e}_j$. Bold face numbers indicate that the constraint is binding.

Table 3.4 reports performances of the equilibrium carbon markets in terms of global welfare and mitigation levels. From columns 4 and 5 of Table 3.4, we find that global welfare (NPV over 100 years) and mitigation of the carbon market are depending on the constraint. Figure 3.1a provides a scatter plot of the global NPV over 100 years achieved by the equilibrium carbon markets for different levels of the constraint. Note that at $\alpha = 0.88$, 0.92 and 0.94 we find multiple equilibrium carbon markets. The trend line in Figure 3.1a shows that the global NPV of the equilibrium carbon markets generally increases when the constraint becomes stricter. Similarly mitigation levels increase as the constraint becomes stricter. Turning to coalition structures we observe that when tightening the constraint '*broad but shallow*' agreements are replaced by '*narrow but deep*' ones. As the constraint becomes stricter it becomes binding for more traders and the shadow value of allowances increases. This indicates increasing forgone payoffs for individual coalition members. At the same time tougher abatement targets of the coalition increase the free-rider incentives and only a small coalition {JPN, RUS} remains to be stable when the constraint tightens. Enhanced welfare effects resulting from more stringent policies can be explained as follows. Carbon sellers, due to the constraints on total emission allowances, have improved payoffs through an increased carbon price and abatement benefits which outweigh larger abatement costs. Carbon buyers who usually obtain higher marginal benefits of global abatement, can gain from the increased mitigation level. Lastly, singletons have higher payoffs from increased global mitigation. Figure 3.1b shows the carbon price in equilibrium carbon markets under different constraints. We can see that the carbon price is generally higher when the constraint becomes tighter. Multiple prices that can be observed at $\alpha = 0.88, 0.92$ and 0.94 relate to the multiple equilibrium carbon markets (e.g. {USA, JPN, ROE, RUS, IND}, {USA, RUS, IND} and {JPN, IND} at $\alpha = 0.88$).

The first row of Table 3.4 shows the best-performing stable carbon market is formed by {USA, JPN, ROE, RUS, IND} under α=0.88. This market has the largest membership and the highest mitigation level. A further tightening of the constraint would further increase mitigation and welfare for the given coalition, however, it will destabilise the coalition since free-rider incentives are stronger under a stricter constraint.

Table 3.4 Performances of equilibrium carbon markets with constrained allowance choices

Value of constraint parameter α	Carbon markets with full stability	Price of emission allowance in 2011 ($/ton CO_2)	NPV of Global payoffs over 100 years (billion $)	Global mitigation in 2011 in % of BAU emissions
0.88	{USA, JPN, ROE, RUS, IND}	6.07	9038.73	6.98
0.74	{JPN, RUS}	22.29	8987.14	6.76
0.76	{JPN, RUS}	19.36	8922.64	6.60
0.78	{JPN, RUS}	16.60	8851.23	6.44
0.80	{JPN, RUS}	14.00	8773.37	6.28
0.88	{USA, RUS, IND}	5.91	8760.03	6.49
0.82	{JPN, RUS}	11.58	8689.66	6.11
0.84	{JPN, RUS}	9.34	8600.66	5.95
0.88	{JPN, IND}	1.23	8576.15	5.48
0.92	{JPN, ROE}	3.24	8540.04	5.62
0.94	{JPN, ROE}	1.92	8443.13	5.49
0.90	{USA, JPN, ROE, RUS, IND}	4.17	8442.83	6.29
0.98	{ROE, HIA, MES}	0.47	8353.10	5.39
0.92	{USA, JPN, ROE, IND}	3.15	8057.79	5.80
0.94	{JPN, ROE, IND, ROW}	0.83	7928.85	5.34
0.94	{USA, JPN, ROE, IND}	1.52	7503.21	5.22

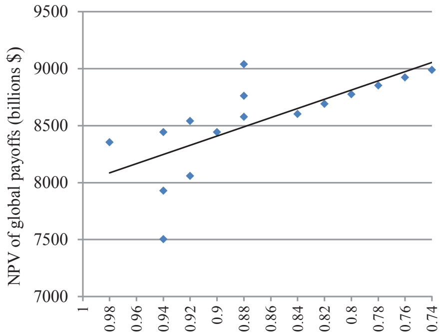

Figure 3.1a. Global NPV of the equilibrium carbon markets under different constraints

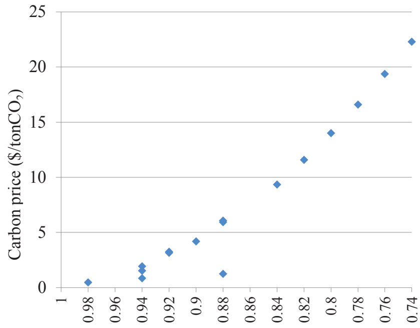

Figure 3.1b. Carbon price in the equilibrium carbon markets under different constraints

In order to gain better insights into the performance of the equilibrium carbon markets under constraints, we report more details of the best-performing carbon market {USA, JPN, ROE, RUS, IND} under $\alpha = 0.88$ in Table 3.5. Column 2 shows that all equilibrium allowance choices are bound by the constraint except for JPN. This is because JPN has the highest marginal benefits from global abatement such that JPN prefers to maintain a high global abatement. This result can also be explained by the shadow value of allowances shown in column 4 of Table 3.5, which indicate that JPN would not lose from a marginal tightening of the allowance constraint. These indicate for all traders except JPN that relaxing the constraint for an individual trader will benefit that trader. Concerning the after-trade abatement the carbon buyers USA and JPN, having the highest MAC, will abate less than their non-cooperative levels (shown in Appendix 3.A2) since they can buy cheap allowances. However the sellers (i.e. ROE, RUS and IND) having lower MAC almost double abatement compared with their non-cooperative abatement levels. The singletons choose the same abatement levels as in the no-trade Nash equilibrium, since they have dominant strategies in the abatement game. Due to increased abatement through carbon trade, the payoffs of all regions are improved as compared to the Nash payoffs (shown in the Appendix, Table 3.A2).

Table 3.5 The best-performing equilibrium carbon market with constraint α=0.88: {USA, JPN, ROE, RUS, IND}

Market participants	Allowance choices in 2011 in % of BAU emissions	Price of emission allowance in 2011 ($/ton$CO_2$)	Shadow value of allowances in 2011 ($/ton CO_2)	Abatement in 2011 (Mton CO_2)	NPV of regional payoffs over 100 years (billion $)
USA	0.88	6.07	3.41	266.89	1808.77
JPN	0.86	6.07	0	47.34	2267.01
ROE	0.88	6.07	5.12	280.04	89.89
RUS	0.88	6.07	2.70	412.07	766.21
IND	0.88	6.07	2.73	852.35	110.75
Singletons					
EUR	----	----	----	403.16	2616.49
OHI	----	----	----	17.80	244.08
HIA	----	----	----	16.29	245.33
CHN	----	----	----	670.99	148.79
MES	----	----	----	9.28	178.61
BRA	----	----	----	8.48	129.26
ROW	----	----	----	142.86	432.66

Overall, our numerical results confirm the cost effectiveness and environmental effectiveness of carbon trade, compared to the no-trade Nash equilibrium. However, in a carbon market with unconstrained allowance choices, the incentive of earning revenues from carbon sales and the arbitrary allowance choices result in a hot-air effect and the external instability of a carbon market with open membership. When imposing constraints on allowance choices, the hot-air effect can be eliminated. Most importantly, by curbing the incentives of obtaining revenues from carbon sale through limiting the allowance choices, the external instability can be reduced. We also find that under a carbon market with constrained allowance choices global mitigation and welfare can be improved most when the constraint is moderate.

3.4.4 Results for carbon markets with Nash-emission levels as baseline for allowance choice constraints

In this chapter we examine the impact of the setting of the baseline on the stability and effectiveness of constrained carbon markets. In the following we assume that allowance constrains are based on non-cooperative Nash emissions levels, i.e. emissions in the All singletons case. Numerical results for all stable carbon markets under the constraint $\alpha \in [0,1]$ are reported in Table 3.6. In general, compared to the BAU baseline for the constraint, results indicate the possibility of larger stable coalitions and more effective markets, i.e. higher global abatement levels in equilibrium. Several features of this result are worth to be highlighted. Firstly, a global carbon market can be sustained when $\alpha = 1$. In this grand carbon market where the upper bound of individual allowance choices is the Nash-emissions level, the global mitigation (5.38%) is equal to the All singletons structure (see Column 5 of Table 3.6; compare to Column 3 of Table 3.1). This result stems from the fact that individual allowance choices of all carbon traders are binding. In this case cooperation does not increase abatement but global payoffs are improved through carbon trade. When the constraint is tighter, at $\alpha = 0.98$ and 0.96, partial carbon markets are stable that comprise of 11 and 8 regions, respectively. The highest global payoffs (10163.68 billion $) can be obtained when $\alpha = 0.98$ and the highest global abatement level (7.27%) can be achieved when $\alpha = 0.96$. These results improve upon global welfare and abatement that can be achieved under the BAU baseline.

The enhanced stability and effectiveness of stable carbon markets with Nash baseline compared to the BAU baseline are related to the binding allowance choices of all carbon traders, which are found in all carbon markets from our numerical results. In particular, when $\alpha = 1$, each region can only improve upon its Nash payoff by joining the carbon market. Thus a carbon market with full participation can be sustained. If the constraint is a little tighter, $\alpha = 0.98$, only ROW prefers to take a free rider position. As argued before, tightening the constraint will always decrease the incentive to sign the agreement. Further tightening of the constraints causes more regions to drop out. From Table 3.6, it can be observed the size of stable carbon markets becomes smaller because of the increased free-riding incentives when the constraint becomes stricter. This is a general finding and robust to changes of the baseline.

Table 3.6 Equilibrium carbon markets with Nash-emissions baseline for individual allowance choices constraints

Value of constraint parameter α	Carbon markets with full stability	Price of emission allowance in 2011 (\$/ton CO_2)	NPV of Global payoffs over 100 years (billion \$)	Global mitigation in 2011 in % of BAU emissions
1	All (Grand carbon market)	1.64	8599.58	5.38
0.98	{USA, JPN, EUR, OHI, ROE, RUS, HIA, CHN, IND, MES, BRA}	3.03	10163.68	7.03
0.96	{USA, JPN, EUR, OHI, ROE, RUS, IND, MES}	6.39	10020.94	7.27
0.94	{JPN, ROE}	5.00	8730.03	5.75
0.94	{EUR, ROE, MES}	9.18	9101.10	6.31
0.92	{JPN, ROE}	6.82	8812.26	5.87
0.92	{JPN, RUS}	9.57	8785.24	5.97
0.90	{EUR, ROE}	14.97	9191.55	6.70
0.90	{JPN, RUS}	11.62	8857.77	6.72
0.88	{EUR, ROE}	17.41	9302.74	6.97

3.5 Conclusions

In this paper, we focus on the conditions for developing a stable international carbon market. Without constraints on individual allowance choices a carbon market can suffer from hot-air effects and market instability. To solve this problem, we consider the role of setting a constraint on allowance choices. Our main findings are the following.

First, under a carbon market with unconstrained allowance choices we find that no stable market emerges. Unconstrained allowance choices can cause hot-air and thus undermine effectiveness and stability of a carbon market.

Second, the stability and the membership of an international carbon market can be increased by imposing a constraint on allowance choices. The reason is that the constrained allowance choices discourage excessive participation of potential sellers. This reduces or avoids hot air. Due to a constraint on the allowance choices, a higher global abatement level can be obtained. Generally, compared to an unconstrained market, constraints can improve the stability and enlarge the scale of an international carbon market, but only to a limited degree. The largest part of the potential gains from cooperation remains unexploited.

Third, when tightening the constraint '*broad but shallow*' agreements are replaced by '*narrow but deep*' ones. In our setting with a constraint on BAU emissions the carbon market with the largest membership can be formed under a relatively lax constraint (12% below BAU emissions in the STACO calibration). When the constraint is closer to the globally optimal abatement, we observe a narrower but deeper stable market with similar performance in terms of global abatement and welfare. We also find that under lax constraints, carbon buyers generally choose their allowances strictly lower than their constraint while sellers choose the binding level. Stricter constraints are binding for all traders. This result is not surprising since the strategic allowance choices by carbon buyers are motivated by the benefits from abatement, while carbon sellers are motivated mainly by revenues of carbon sales. These different motivations induce the strategic allowance choices to depend on the strictness of constraints.

Fourth, our results also point at an alternative option for stabilising an international carbon market. As external instability is an issue, limiting access will increase abatement and global welfare. In fact limiting access can be more effective than allowance choice constraints based on BAU emissions; compare the best performing markets in Tables 3.1 and 3.4. This is in line with the conclusion by Finus (2008) that an open membership regime adopted in current international climate negotiations should be critically reviewed. However this conclusion does not carry over to the case of constraints based on Nash emissions.

Finally, we demonstrate that by tying individual allowance choice constraints to the Nash-emissions levels, a carbon market with full participation can be sustained when the constraint is sufficiently lax (i.e. α close to 1). Different from BAU-related baselines, Nash-related baselines are always binding. Our result indicates that a Nash-related baseline is more successful in terms of global welfare and abatement, as it responds better to individual incentives to participate. In the current policy debate BAU emissions are still dominant for defining and negotiating abatement targets or emission allowances. Our finding suggests that a revision of the baseline could ease negotiations.

A limitation of our analysis is that the constraint on allowance choices is modelled as an exogenous parameter which is common for all carbon traders. A valuable extension of the model would be to include a pre-negotiation stage where the set of all regions first determines the constraint – possibly conditional on regional characteristics – and only after that the membership and allowance choices would follow.

Acknowledgments We thank two anonymous reviewers for helpful comments. We are grateful to Rob Dellink and Miyuki Nagashima who have supported our numerical analysis.

The first author acknowledges support by the Erasmus Mundus EURASIA 2 (www.eurasia2.cz) programme.

3.6 Appendix

Table 3.A1. Benefits share, marginal benefits and BAU emissions in the STACO model

Regions	Benefits share of global abatement	Marginal benefits in 2011 ($/ton CO_2)	Uncontrolled Emissions (BAU) in 2011 (Gton CO_2)
USA	0.2265	6.49	7084.96
JPN	0.1725	7.34	1386.55
EUR	0.2360	8.92	4891.74
OHI	0.0345	0.80	1411.23
ROE	0.0130	0.30	1503.11
RUS	0.0675	2.43	2234.26
HIA	0.0300	0.81	2591.49
CHN	0.0620	0.57	11623.98
IND	0.0500	0.34	3276.20
MES	0.0249	0.60	991.67
BRA	0.0153	0.43	2181.62
ROW	0.0680	1.49	5660.27
Sum	1.0000	------	44837.08

Table 3.A2. Performances of the non-cooperative Nash equilibrium and the first-best social optimum scenario

Regions	Nash equilibrium		Social optimum	
	Abatement in 2011 (Mton CO_2)	NPV of payoffs over 100 years (billion $)	Abatement in 2011(Mton CO_2)	NPV of payoffs over 100 years(billion $)
USA	284.63	1742.81	1240.94	6558.90
JPN	56.97	2125.91	222.85	9310.70
EUR	403.16	2428.74	1130.12	10128.17
OHI	17.80	228.31	390.08	707.83
ROE	71.36	85.83	614.66	19.26
RUS	266.34	686.37	904.47	2421.62
HIA	16.29	228.75	531.82	345.23
CHN	670.99	138.94	3740.05	-727.20
IND	465.01	90.61	1586.02	0.88
MES	9.28	166.47	470.65	92.81
BRA	8.48	120.49	330.59	193.56
ROW	142.86	403.16	1728.81	507.82
Global	2413.17	8446.39	12891.06	29559.58

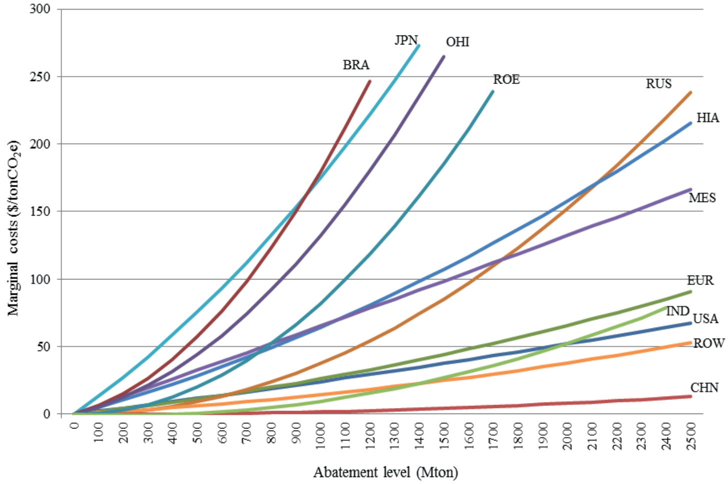

Figure 3.A1. Marginal abatement cost curves in 2011 in the STACO model

Chapter 4

Nash bargaining solutions for international climate agreements under different sets of bargaining weights*

Bargaining is a tool to share collaborative gains and to facilitate reaching an agreement. To improve incentives to join an international climate agreement (ICA), the Nash bargaining solution can be used to distribute cooperative gains across signatories. In this paper, we examine how the formation of ICAs and their mitigation efficiency are impacted by the use of the Nash bargaining solution. In a Nash bargaining game with heterogeneous players, bargaining powers are unequal and may be driven by different characteristics of the players. We employ different sets of asymmetric bargaining weights in order to examine the effectiveness of climate coalitions that emerge as stable agreements. Using the Nash bargaining solution, we obtain results from the Stability of Coalition model (STACO). We find that the Nash bargaining solution can improve the participation incentives and performances of ICAs as compared to agreements that do not redistribute gains from cooperation, but its capacity to overcome free-riding incentives is limited. However, if Nash bargaining accounts for outside options of players, we find larger stable coalitions and higher global abatement levels. In fact, Nash bargaining with outside options can stabilise the largest coalitions that can possibly be stable in our game.

* Published in Journal of International Environmental Agreements: Politics, Law and Economics:
Yu S., van Ierland E.C., Weikard H.-P., Zhu X. (2017). Nash bargaining solutions for international climate agreements under different sets of bargaining weights, DOI: 10.1007/s10784-017-9351-3.

4.1 Introduction

Mitigation of greenhouse gas (GHG) emissions is costly. Due to the public-good property of GHG emissions mitigation each individual country has an incentive to free-ride on the abatement efforts of other countries. However, through multilateral negotiations an international climate agreement can be formed to alleviate the social dilemma (Carraro and Siniscalco 1993; Finus 2003). A climate agreement comprises a mitigation target for members, but also needs to distribute gains from cooperation.

The distribution of coalitional gains across countries can be organised through transfer schemes which are effective instruments to offset free-riding incentives and improve the stability of International Climate Agreements (ICAs) (Carraro and Siniscalco 1993; Barrett 1997; Botteon and Carraro 1997; Rose et al. 1998; Rose and Stevens 1998; Weikard et al. 2006; Carraro et al. 2006; Weikard 2009; Nagashima et al. 2009). Sharing the gains of cooperation is solving a bargaining problem (Nash 1950; Powell 2002), hence the Nash bargaining solution (NBS) can be used to determine transfer schemes for ICAs. Carraro and Siniscalco (1993) analyse the role of welfare transfers for coalition stability among symmetric players when transfers are determined by applying the Nash bargaining solution. Their theoretical results show that the size of stable coalitions can be extended by 'bribing' singletons with transfers. Botteon and Carraro (1997) and Carraro et al. (2006) apply the Nash bargaining rule to surplus sharing within coalitions. They extend the analysis to asymmetric players and their results are derived resorting to numerical analysis. However, previous studies on the application of a Nash bargaining rule assume equal bargaining weights among countries and do not take the different bargaining powers of asymmetric countries into account. Bargaining powers determine the sharing of collective gains which, in turn, determines the incentives to join a climate agreement. In this paper, we assume that the distribution of gains is the outcome of a bargaining process with unequal bargaining power and, thus, transfers are determined by bargaining power. We employ the Nash bargaining solution to model the distribution of cooperative gains. The bargaining outcome is subject to bargaining power. The key issue is what constitutes the bargaining power of each negotiator. Costantini et al. (2015) provide a sketch of future potential bargaining positions of developing countries in climate negotiations according to key structural features like economic power, geographic, environmental and social characteristics, and the energy system. Considering the importance of bargaining power in climate negotiations, we discuss and review potential reasons that could induce differences of negotiators' bargaining power in international climate negotiations. Thus determinants of bargaining power can be identified and then used for the quantification of negotiators' bargaining weights. Furthermore, we model the international climate negotiations as a Nash bargaining game in which cooperative gains are distributed based on the NBS with asymmetric bargaining power, and we study which climate coalitions could form, given the different bargaining powers of the negotiators.

Although Nash bargaining has been used as a transfer scheme in the literature on ICAs, it has not been investigated in depth for the surplus sharing of coalitional gains among countries with unequal bargaining powers. The novelty of this paper is that we explore the impact of using the NBS for distributing coalitional gains under different sets of bargaining weights on the stability and effectiveness of international climate agreements. We consider different possible determinants of bargaining power which are exogenously determined in our model. Our analysis complements the set of coalitional surplus sharing rules.

In this paper, we identify five different factors that can determine negotiators' bargaining power. (i) In bargaining theory, time preference, i.e. the willingness to wait for the payoff is often proposed as a driving factor that can influence the distribution of gains. A player who is more patient has more bargaining power (Rubinstein 1982; Binmore et al. 1992; Powell 2002). Thus we use time preference represented by the discount factor to determine bargaining power of negotiating countries. (ii) In climate negotiations, it is commonly argued that the distribution of cooperative payoffs should be in accordance with abatement efforts. Larger efforts give a claim to a larger share of the gains. Hence, we use the proportion of individual abatement in a coalition to represent each negotiator's bargaining weight. (iii) Another way to assess efforts is to use abatement costs. This is different from (ii) if countries differ in marginal costs of abatement. Larger costs could justify a larger claim. Thus we use countries' total abatement costs, reflecting their monetary efforts as an indicator of bargaining power. (iv) Abatement benefits of a country reflect the avoided damages from reducing GHG emissions which are associated with a country's vulnerability to climate change. A country that is more vulnerable to climate change will be more eager to get involved in the climate cooperation than a country that is less vulnerable. Hence, we use the inverse of abatement benefits as an indicator of bargaining power. (v) In international negotiations, economically powerful countries can be more successful in shaping the agreement. We therefore take the economic power measured by gross domestic product (GDP) as an indicator of bargaining power.

To compare and examine the impact of the NBS with different sets of unequal bargaining weights on incentives to cooperate, we formulate a two-stage cartel formation game. At the first stage, each country decides whether to participate in a climate coalition or not by evaluating payoffs received from being a signatory or a singleton. At the second stage, abatement targets are set cooperatively by coalition members, but their individual payoffs are determined by applying the NBS with a given set of bargaining weights.

A bargaining solution is not just affected by bargaining power, but also by players' outside options (Wagner 1988; Powell 2002). A decision to take up an outside option implies a withdrawal from the bargaining process. The outside option payoff imposes a lower bound on the bargaining solution since no one must accept an agreement that makes him worse off compared to what he can obtain otherwise (Binmore et al. 1992; Muthoo 1999). For the bargaining problem of international climate agreements, we assume that a country's outside option is to abstain from an agreement and to remain a singleton player. This is in line with Muthoo (1999, p. 105) who explains that outside options do not affect the disagreement point. The Nash bargaining solution with outside options falls in the class of optimal sharing rules described by Carraro et al. (2006) and Weikard (2009). Hence, our paper provides an additional motivation for the use of optimal sharing rules.

We adopt the concept of internal and external stability to analyse our game (cf. D'Aspremont et al. 1983). An agreement is internally stable if no member wants to leave. It is externally stable if no other player wants to join. This implies that we consider only single deviations which define the outside option payoffs. If a deviation would trigger others' withdrawal from the coalition, a simple internal/external stability concept would not be adequate and more sophisticated solution concepts such as coalition proof equilibrium (Bernheim et al. 1987) or farsighted stability (Chwe 1994) could be employed. De Zeeuw (2008) shows that farsighted coalition stability can lead to larger stable coalitions with higher global welfare. Another assumption that we adopt is that we allow only for one coalition. A deviator from the coalition becomes a singleton and cannot make an agreement with other players. Allowing for multiple coalitions would lead to higher abatement levels and global welfare as has been shown by Asheim et al. (2006) and Sáiz et al. (2006). However, in this paper we do not consider the possibility of multiple coalition structures and we confine the analysis to internal and external stability. The implications of refined solution concepts is left to future research.

To see how different sets of bargaining weights and the relevance of outside options impact coalition stability, we compare results from the STACO (Stability of Coalitions) model. STACO is a global model with calibrated abatement costs and benefits functions for 12 world regions. We use it to test stability of the $2^{12} - 12$ possible coalitions that may form. Our results provide implications of sharing rules based on the NBS, and also its impact on the formation and stability of ICAs.

In what follows, we present the game in Section 4.2. Section 4.3 discusses potential determinants of bargaining power, i.e. we introduce different sets of bargaining weights used in the NBS for distributing cooperative gains. Section 4.4 describes the STACO model and our numerical results in more detail. The paper ends with discussions and conclusions in Section 4.5.

4.2 The game theoretical model

We consider a set of players $N = \{1,2,\ldots,n\}$ representing countries or regions that negotiate an agreement on mitigating GHG emissions. We allow for asymmetric abatement benefits and costs. The formation of a climate agreement is modelled by the following a two-stage game.

Stage 1: All players ($i \in N$) announce whether to sign an agreement or not. Their decisions are made non-cooperatively and simultaneously. Formally, an agreement is a subset of the set of players. The set of all possible agreements is:

$$\zeta = \{S \subseteq N | \ s \geq 2\}, \qquad |\zeta| = 2^n - (n+1).$$

where $s = |S|$ represents the number of signatories. If negotiations fail, then there is no agreement and all players remain singletons.

Stage 2: At the second stage signatories S and the remaining singletons play a transboundary pollution game. Abatement strategies are chosen simultaneously by signatories and singletons. Signatories ($i \in S$) decide on their abatement by maximising their aggregated payoffs; non-signatories ($i \in N \setminus S$) maximise their individual payoff. If no agreement has been reached at stage 1, payoffs are determined by the non-cooperative equilibrium outcome of a n-player transboundary pollution game. We denote this outcome by D as it represents the disagreement point of the bargaining game.

The set of abatement choices by all players can be defined as $q = (q_1, \ldots, q_n)$ with the condition $q_i \in [0, \bar{e}_i]$, where $\bar{e}_i$ denotes the business-as-usual emissions level. Once abatement is chosen, the individual signatory's payoff denoted by $\pi_i(q)$ can be determined, where q denotes the abatement vector. The coalitional gains of the agreement S are the difference between the aggregate payoff of signatories $\sum_{i\in S}\pi_i(q)$ and what they would get in case of disagreement. The final payoffs of signatories denoted by π_i^* are based on the NBS used to redistribute the coalitional gains. The set of weights reflecting bargaining power is denoted by $\{\lambda_i\}_{i\in S}$. It is exogenously determined in our model.

Note that in our game, we only consider the formation of one single agreement, such that a player deviating from S becomes a non-signatory. Therefore, the outside option for each signatory $i \in S$ in our Nash bargaining game is the payoff received as a singleton when other players maintain their membership status. In this setting, for an agreement S with only two signatories (i.e. $s = 2$), the outside option is identical to the disagreement point or status quo, but in general this is not the case. Obviously, an agreement S can be accepted by a signatory if and only if its coalitional payoff is larger than its outside option payoff.

We solve this two-stage game by backward induction in order to identify the sub-game perfect Nash equilibria. At stage 2, the signatories $i \in S$ choose their optimal abatement levels q_i by maximising joint payoffs denoted by $\sum_{i\in S}\pi_i$. We have the following maximization problem for signatories $i \in$ S:

$$\max_{q_i} \quad \sum_{i\in S}\pi_i(q) = \sum_{i\in S}\big(B_i(q) - C_i(q_i)\big). \tag{4.1}$$

where $q = \sum_{i\in N} q_i$ denotes the global abatement level. The abatement cost function denoted by $C_i(q_i)$ is increasing and strictly convex, i.e. $C_i'(q_i) > 0$ and $C_i''(q_i) > 0$. Abatement benefits $B_i(q)$, depend on the overall level of abatement q and are assumed to be linearly increasing, i.e. $B_i'(q) > 0$ and $B_i''(q) = 0$, implying a dominant strategy for each player. The equilibrium condition for signatories $i \in S$ is derived from the first order condition of problem (4.1):

$$C_i'(q_i^*) = \sum_{j\in S} B_j'(q^*). \tag{4.2}$$

At this stage, each non-signatory maximises its own payoffs. The problem for non-signatories can be formulated as follows:

$$\max_{q_i} \quad \pi_i(q) = B_i(q) - C_i(q_i). \tag{4.3}$$

The equilibrium condition for singletons $i \in N \backslash S$ can be obtained by deriving the first order condition of the problem (4.3). This gives $B_i'(q^*) = C_i'(q_i^*)$.

Based on abatement choices of all players, the payoffs of signatories and singletons can be determined. Signatories redistribute the aggregated payoffs $\sum_{i \in S} \pi_i(q)$ based on the NBS. Their bargaining powers are unequal and given by a set of bargaining weights $\{\lambda_i\}_{i \in S}$. The redistributed payoff under the NBS can be represented by a set denoted by $\{\pi_i^*\}_{i \in S}$, which solves the Nash bargaining problem described as follows:

$$\max_{\{\pi_i^*\}_{i \in S}} \prod_{i \in S} (\pi_i^* - \bar{\pi}_i)^{\lambda_i}, \tag{4.4}$$

s.t.

$$\pi_i^* \geq \bar{\pi}_i,$$

$$\pi_1^* + \cdots + \pi_S^* = \sum_{i \in S} \pi_i(q).$$

in which $\bar{\pi}_i$ represents the non-cooperative payoffs of signatories $i \in S$, which is the disagreement point ($(\bar{\pi}_i)_{i \in S} = D$). We assume that gains from cooperation can be redistributed among signatories without incurring transactions costs. Therefore, the bargaining set is convex and compact which ensures the uniqueness of the solution of bargaining problem (4.4). Signatories' bargaining weights satisfy the condition $\sum_{i \in S} \lambda_i = 1$. A higher value of λ_i indicates a strategic advantage in the bargaining process.

At stage 1, all players decide to sign an agreement or not. Here, we use the partition function $V_i(S)$ that can be derived from the solution of the stage-2 game to represent each player's payoffs under the coalition S based on the NBS. Note that a signatory receives its outside option payoffs when deviating from the coalition S, denoted by $V_i(S \backslash \{i\})$. The Nash equilibrium of the stage-1 game satisfies the following internal and external stability conditions (d'Aspremont et al. 1983).

Internal stability:

$$V_i(S) \geq V_i(S \backslash \{i\}), i \in S. \tag{4.5}$$

External stability:

$$V_i(S) \geq V_i(S \cup \{i\}), i \in N \backslash S. \tag{4.6}$$

4.3 The representation and interpretation of different sets of bargaining weights

In this section, we discuss factors that could lead to differences in countries' bargaining power in international climate negotiations. Based on these factors, weights of bargaining power can be identified that will be used to determine the bargaining outcomes and, hence, the individual payoffs for all coalition members. We also explain the relevance of outside options in our game and discuss their role for stable climate agreements.

4.3.1 Bargaining weights based on discount factor

Gains from cooperation can only be obtained when agreement is reached. A delay of reaching an agreement is costly (Rubinstein 1982; Muthoo 1999). This is particularly relevant for climate agreements: the sooner cooperation is achieved, the smaller the climate damages will be (Courtois and Tazdaït 2014). Binmore et al. (1986) show that in a bargaining game players' time preferences impact their strategic choices and thus the bargaining solution. The discount factor reflects a player's willingness to wait and can be used as an indicator of the negotiators' bargaining power.

Over time the net present value (NPV) of the gains from cooperation falls quicker for a region with a higher discount rate than for a region with a lower discount rate. Therefore, the higher a region's discount rate, the stronger its preference for an ICA that is formed and implemented earlier. Regions with a strong preference to reach an agreement are more willing to give in. They will have a larger cost of waiting and therefore less bargaining power.

Let r_i be the discount rate prevailing in region i. We use the discount factor denoted by $\delta_i \equiv \frac{1}{1+r_i}$ to represent the bargaining power of signatory i in the negotiation. The corresponding bargaining weight can be represented as:

$$\lambda_i = \frac{\delta_i}{\sum_{j \in S} \delta_j}, i \in S. \tag{4.7}$$

Under this set of bargaining weights, regions with relative lower discount rates are expected to have stronger participation incentives induced by higher bargaining weights and a larger share of coalitional gains.

4.3.2 Bargaining weights based on abatement efforts

As compared to developed regions with high marginal abatement costs and benefits, developing regions with low marginal abatement costs and benefits have less incentives to join a climate coalition. The main reason is that developing regions contribute larger shares of global abatement but with lower private returns. Considering the importance of developing regions' contribution to abatement, it can be argued that regions engaging in greater abatement efforts, if they join, can ask for a larger share of the gains. Hence, we use the proportion of individual abatement in a coalition to represent each negotiator's bargaining weight. The larger the mitigation efforts of a coalition member, the larger its bargaining power in the negotiation over coalitional gains. The bargaining weight based on abatement efforts can be formulated as:

$$\lambda_i = \frac{q_i^*}{\sum_{j \in S} q_j^*}, i \in S. \tag{4.8}$$

in which, q_i^* is the equilibrium abatement level of each signatory i of coalition S. It can be expected that under this set of bargaining weights, regions contributing larger shares to the coalitional abatement will have more incentives to participate.

4.3.3 Bargaining weights based on abatement costs

Within a climate coalition, due to differences in the technological development and the resulting differences in marginal abatement costs between countries, a large abatement assignment of a member does not necessarily imply high total abatement costs if marginal abatement costs are low. Hence, abatement effort q_i does not accurately reflect the cost each member pays for cooperation. The controversy about collaborative gains allocation that is induced by countries' uneven costs for mitigation cooperation has been a recurring issue put on the negotiation table (Barrett and Stavins 2003; Carraro et al. 2006). Generally, countries taking on larger abatement costs would claim compensation in the form of a larger share of the cooperative gains. If the total abatement is seen as a joint effort requiring investment, then the gains from cooperation should be distributed proportional to these investments. Hence, a country's abatement cost can be identified as a source of asymmetry in bargaining power during negotiations on climate cooperation. We take the abatement cost of a country as a measure of its bargaining power. The bargaining weight can be formulated as follows:

$$\lambda_i = \frac{C_i(q_i^*)}{\sum_{j \in S} C_j(q_j^*)}, i \in S. \tag{4.9}$$

Under such a set of bargaining weights, the higher a country's abatement cost, the higher will be its share of the gains from cooperation.

4.3.4 Bargaining weights based on climate change damages

Damages resulting from climate change differ across regions due to different impacts of climate change, different economic losses and different valuations of environmental quality. Regions facing high damages are eager to mitigate climate change. However the low-damage regions have less incentives to join the cooperation. This difference in preferences for climate cooperation implies that low-damage countries hold more bargaining power than high-damage countries.

In our model abatement benefits are avoided climate change damages. Hence, we use abatement benefits to represent climate change vulnerability. Bargaining power is then inversely related to climate change vulnerability. The bargaining weight based on damages can be represented as

$$\lambda_i = \frac{\frac{1}{B_i(q^*)}}{\sum_{j \in S} \frac{1}{B_j(q^*)}}, i \in S. \tag{4.10}$$

From Eq. (4.10) it is straightforward to see that higher benefits from global abatement are associated with lower bargaining power and a lower bargaining weight λ_i.

4.3.5 Bargaining weights based on economic power

In international negotiations among asymmetric regions, the economic power reflected by the GDP of a region can affect its bargaining power. Generally, regions which are characterised by a larger GDP can exert more influence over the regions with a lower GDP, for example through issue linkage. Issue linkage connects environmental negotiations with other economic issues (e.g. trade or technological cooperation). Issue linkage can offset free-riding incentives in climate cooperation (Carraro and Siniscalco 1993, 1995). In climate negotiations, regions with an economic advantage can put pressure on poorer regions with 'sticks and carrots'. For example, they can withhold or offer technological cooperation. Therefore, regions who have advantage in economic power have larger bargaining power in climate negotiations. Based on this argument, economically powerful regions can obtain a larger share of the joint payoff. This reasoning has also been put forward by Rose et al. (1998) where they discuss a transfer rule based on income claims. Determining negotiators' bargaining power by economic power, the bargaining weight can be represented as:

$$\lambda_i = \frac{GDP_i}{\sum_{j \in S} GDP_j}, i \in S. \quad (4.11)$$

4.3.6 Outside options

In the bargaining game outside options introduce a lower bound on each player's payoff. The presence of outside options can thus affect the negotiation outcome by considering the impact of the minimum payoff a player can assure for himself when leaving the negotiation. Hence, in our game we assume that a region's outside option is to abstain from an agreement and to remain a singleton player. As mentioned before, a player's outside option payoff can be written as $V_i(S \backslash \{i\})$. In the literature examining coalition stability in cartel games it has been pointed out that coalition S can be internally stable whenever the coalition payoff does not fall short of the sum of members' outside options, i.e.

$$\sum_{i \in S} V_i(S) \geq \sum_{i \in S} V_i(S \backslash \{i\}). \quad (4.12)$$

Inequality (4.12) is a necessary condition for internal stability of coalition S. If we consider a bargaining solution with outside options, then the outside option payoff is guaranteed for each signatory whenever (4.12) is satisfied. Carraro et al. (2006) and Weikard (2009) have called solutions satisfying internal stability when (4.12) holds "optimal transfers", since payoffs are re-arranged to incentivise participation in the agreement. It is obvious then that bargaining with outside options belongs to the class of optimal transfer rules. It is important to note that the set of stable coalitions is not affected by the distribution of the surplus that remains after all signatories have received their outside option payoffs. This implies that in our game, under bargaining with outside options, the set of stable coalitions is independent of bargaining weights λ_i. Hence, for the analysis of stability there is no need to apply different sets of bargaining weights.

4.4 Simulations and Results

In this section, we describe the implementation of the simulation analysis employing the STACO model in Section 4.4.1. Section 4.4.2 presents the simulation results and a discussion of our findings concerning stability and effectiveness of coalitions under the NBS. We examine the relevance of outside options in Section 4.4.3. All results and discussions in this section are based on the STACO 3.0 model as documented by Nagashima et al. (2011) and Dellink et al. (2015).

4.4.1 Simulations employing the STACO model

To gain practical insights into the implications of the NBS with different sets of bargaining weights, we employ a numerical simulation model, the STACO model. The STACO model is a combined game-theoretic and integrated assessment model created to examine the formation and performances of international climate agreements among twelve world regions as specified in the Table 4.A1 in the Appendix (cf. also Nagashima et al. 2011; Dellink et al. 2015). By specifying the abatement benefit and cost functions for these twelve heterogeneous regions, the STACO model enables us to analyse coalition stability, abatement levels, efficiency and welfare. Considering inertia of the climate system, the simulation analysis in STACO adopts a 100-year time horizon (from 2010 to 2110). The formation of an ICA in STACO is a standard two-stage cartel formation game, in which the membership decision is taken once and for all periods. The abatement choice of each player i for the whole time horizon can be represented by an abatement path $(q^*_{i,1},..., q^*_{i,100})$. Accordingly, the payoffs of each player over the planning horizon can be represented as a payoff path $(\pi^*_{i,1}, \dots, \pi^*_{i,100})$. Finally, the stability of a climate coalition is checked based on the net present value of the payoff stream over 100 years according to the internal and external stability conditions (4.5) and (4.6). For a full specification of the latest version of the STACO model (STACO 3.0) the reader is referred to Nagashima et al. (2011) and Dellink et al. (2015).

For our simulations we use the numerical computing software MATLAB. For each period, signatories and singletons decide their optimal abatement $q^*_{i,t}$ according to Eqs. (4.2) and (4.3). Accordingly, the payoffs of signatories can be derived in each period based on Nash bargaining solutions described in Eq. (4.4). The payoffs of singletons are the net gains from abatement. In particular, each signatory's bargaining weight $\lambda_{i,t}$ in each period is based on the value of different determining factors at each period, for example, the discount factor $\delta_{i,t}$, abatement $q^*_{i,t}$, abatement costs $C_{i,t}(q^*_{i,t})$, climate change damages $\frac{1}{B_{i,t}(q^*_t)}$ and economic power $GDP_{i,t}$. For each set of bargaining weights, we calculate the abatement and payoffs path of each region under every coalition that can be formed. Finally, based on the calculation of the NPV of each player's payoff stream over 100 years we perform a stability check for each coalition according to Eqs. (4.5) and (4.6).

4.4.2 Results for the Nash bargaining solution without consideration of outside options

In this section, we report and discuss results on coalition formation and performances under the NBS with different sets of bargaining weights. The role of outside options is discussed in Section 4.4.3.

To compare our surplus sharing scheme based on the NBS with other conventional sharing schemes that are frequently discussed (for example egalitarian, historical responsibility or ability to pay) we calculate transfers generated in NBS under the five different sets of bargaining weights for the grand coalition. Table 4.1 illustrates the amount of transfers for each region under various sets of bargaining weights. Regions with negative transfers are payers, and regions with positive transfers are receivers. Due to high marginal abatement benefits USA, JPN and EUR can benefit more than other regions. Hence, USA, JPN and EUR are always transfer payers under all five sets of bargaining weights. This result is in line with the result under the egalitarian and ability-to-pay sharing scheme calculated by Altamirano-Cabrera and Finus (2006) using STACO. However, JPN becomes transfer receiver under a historical responsibility rule. This is due to the low Business-As-Usual emissions of JPN, implying that JPN contributes less to the current GHG concentration and thus JPN has a low mitigation target. In terms of the amount of transfers JPN and EUR are the two largest payers and their payments are much larger than under the egalitarian and ability-to-pay rule in Altamirano-Cabrera and Finus (2006). In contrast, USA pays the largest amount of transfers under the egalitarian and ability-to-pay rule in Altamirano-Cabrera and Finus (2006). This is due to USA's relatively large population and high GDP per capita. From Table 4.1, it can also be seen that CHN and BRA always receive the largest amounts of transfers under different bargaining weights. This also explains why CHN and BRA have strong participation incentives as shown in Table 4.2. When other conventional sharing rules are applied, IND and BRA are the largest transfer receivers (Altamirano-Cabrera and Finus 2006).

Now we turn to the stability results when the NBS is used to redistribute coalitional gains. Firstly, our results show that all coalitions with two members are internally stable. As explained by Weikard et al. (2006) this result is related to the linear functional form of abatement benefits and non-negative weights, which ensure a positive coalitional surplus. Thus being a signatory of a two-members coalition gives a larger payoff than that of being a singleton. Table 4.2 shows the results for stable coalitions and the value of bargaining weights for each signatory under different determinants of the bargaining power. The second column of Table 4.2 lists all stable coalitions for the five sets of asymmetric bargaining weights described in Section 4.3 and a reference scenario with equal weights. Under the scenario with equal bargaining weights, three coalitions (i.e. {RUS, CHN, BRA}, {HIA, CHN, BRA} and {CHN, IND, BRA}) are stable. In STACO, regions like RUS, CHN and IND are characterised by flat marginal abatement cost curves and by a moderate level of abatement benefits (see Table 4.A2 in Appendix 4.6 and Figure 3.A1 in Appendix 3.6). In a coalition of regions with similar marginal abatement costs and benefits, participation incentives can be promoted and maintained with equal sharing of collective gains. Regions BRA and HIA face steep marginal abatement cost curves and low marginal benefit shares. Hence when joining a coalition the required additional abatement remains limited. It is CHN that undertakes the largest abatement efforts but also not much more than under All Singletons.

As shown in Table 4.2, generally, there are small stable coalitions under all sets of bargaining weights we examine. The reason is that for each set of asymmetric bargaining weights, only the regions with an advantage have sufficient incentives to join. For example, as our numerical results show, USA has strong incentives to join an ICA with members of HIA, CHN and BRA when the bargaining weight is based on GDP, whereas among six stable coalitions only China, India and the Rest of the World are motivated to join with large incentives when the bargaining weight is determined by abatement efforts. Our results also show that multiple equilibrium coalitions emerge under each set of bargaining weights. Generally, among all determinants, more coalitions can be stabilised when bargaining power is determined by signatories' abatement efforts, damages and abatement costs. It should be noticed that same stable coalitions can emerge under different sets of bargaining power, for example, coalitions {OHI, CHN}, {ROE, CHN} and {RUS, CHN} can be stabilised with bargaining weights based on 'abatement efforts' and 'abatement costs'. Even though some regions' bargaining power can be interpreted as different values under different determinants, and the corresponding NBS in terms of the distribution of coalitional gains is also changed accordingly, the same equilibria with respect to the stable coalitions can still be reached. It is also interesting to see that, even though there are multiple equilibrium coalition structures for each distribution of bargaining weights, one coalition member is present in each equilibrium structure (CHN or BRA). This is due to their relatively larger bargaining weights under the respective set of weights (see Table 4.2).

When the bargaining weight is determined by regions' discount factors, regions with high discount rates, for example CHN and IND, cannot strike an agreement with developed regions like USA, JPN and EUR which have low discount rates and induce large additional abatement efforts of CHN. Due to low discount rates (see Appendix A2), developed regions like USA, JPN and EUR are in a better bargaining position compared to regions that have high discount rates and CHN cannot recover the cost of abatement. However, as can be seen from Table 4.2, cooperation can be established between CHN and regions like ROE, BRA and RUS (e.g. {ROE, CHN, BRA}, {RUS, CHN, BRA}). To see why such coalitions can be stable, notice that ROE, RUS and BRA have low marginal benefits, requiring little extra abatement compared to disagreement. To shed more light on stability consider Figure 4.1 which depicts the payoff space for RUS and CHN in a coalition with BRA. This coalition is stable under equal and discounting bargaining weights. In Figure 4.1, $\mathrm{D} = (836.58, 283.52)$ is the disagreement point. BRA's payoff in that point is 244.52 and not reported in the Figure. The downward sloping line depicts any payoff distribution between CHN and RUS when BRA receives its outside option payoff, a minimum requirement for internal stability. Hence to the right of that line where RUS and CHN receive more and BRA receives less the coalition cannot be stable. The dashed vertical and horizontal lines depict the outside option payoffs of RUS and CHN, respectively. Only bargaining solutions that produce payoff vectors in the shaded triangle will be internally stable. We highlight three points in Figure 4.1. Point $\mathrm{B} = (884.80, 289.66)$ is the best payoffs that BRA can obtain (288.22) in a stable coalition with RUS and CHN. Both, RUS and CHN receive only their outside option payoff. Points $\mathrm{E} = (884.92, 295.66)$ and $\mathrm{F} = (885.31, 294.38)$ are the bargaining solutions for equal bargaining weights and weights based on discounting. Both are stable as can be seen from Table 4.2 and very close to each other as can be seen from Figure 4.1.

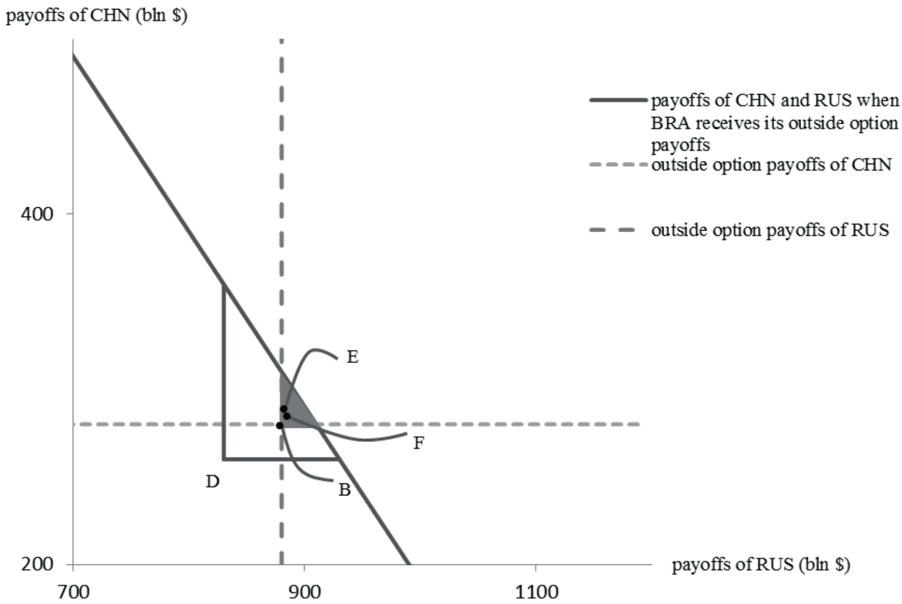

Figure 4.1. The bargaining set for the coalition {RUS, CHN, BRA}. The disagreement point is D. The bargaining set is the triangle to the north-east of D. Point B marks the outside option payoffs for RUS and CHN. The shaded triangle marks internally stable allocations.

In scenarios where bargaining weights are based on abatement efforts and abatement costs, China is a member of all stable coalitions. This can be attributed to China's advantage in terms of low marginal abatement costs, which makes China contribute large shares to the total coalitional abatement. The large contribution to coalitional abatement puts China in a strong bargaining position and it therefore receives a larger share of the gains. However, there are two exceptions of stable coalitions (i.e. {CHN, IND} and {CHN, ROW}), in which the bargaining weights of CHN are lower compared to IND and ROW (i.e. $\lambda_{CHN} = 0.47$ and 0.39), see Table 4.2. This is due to the lower marginal abatement costs of IND and ROW as compared to CHN at the equilibrium abatement level.

Under bargaining weights determined by climate change damages, BRA always has more bargaining power because of its lowest benefit share. Hence, as shown in the scenario for damage weights in Table 4.2, BRA has generally a higher bargaining weight than its respective coalition partner (except when it forms a coalition with India) and thus appears in all stable coalitions in this scenario. Under bargaining weights determined by the economic power, USA holds an advantageous bargaining position in the negotiation. As shown in the last row of Table 4.2, USA has the largest bargaining weight in the coalition {USA, HIA, CHN, BRA}. However, USA would prefer one of the other five bargaining weight scenarios since in these scenarios it would benefit from being a free-rider. Our results also show that a stable cooperation between two regions with equal bargaining weights, like CHN and ROW, can also be reached.

In order to compare the performance of stable coalitions that we find under different sets of bargaining weights, we report more detailed results in Table 4.3. The table shows results for the best-performing stable coalitions in terms of the net present value of global payoffs for each set of bargaining weights. Table 4.3 shows that in general the set of bargaining weights that favours large emitters can lead to higher abatement and welfare levels.

Among the five sets of bargaining weights, the highest global abatement and welfare can be obtained under the coalition {USA, HIA, CHN, BRA} which is stable for bargaining weights determined by GDP. The reasons for this finding are, firstly, that the size of this coalition is the largest among all stable coalitions under different sets of bargaining weights; secondly, this is due to the participation of the world's two biggest GHG emitting countries USA and CHN. Their GHG emissions account for a large part of the world emissions, hence the abatement level adopted by these two countries is also prominent for global abatement and welfare. This result also reflects the important impact of the participation by USA and CHN in the formation of ICAs. By contrast, the coalition {CHN, BRA}, that is stable when the bargaining weights are determined by damages, offers the lowest global abatement and welfare in our set of scenarios. It is even Pareto dominated by {USA, HIA, CHN, BRA}. This result is straightforward to understand: compared to other regions BRA is a region with higher marginal abatement costs. Therefore, the equilibrium abatement level by BRA is lower, which results in the lowest global abatement and welfare obtained by coalition {CHN, BRA}. In the reference scenario with equal bargaining weights, the global abatement of the stable coalition {CHN, IND, BRA} is lower than what is achieved under other sets of asymmetric bargaining weights, except for the weights derived from damages. Nevertheless, the welfare obtained is higher, unless weights are determined by economic power. It can be concluded that signatories with low marginal abatement costs forming a coalition with a region with high marginal benefits lead to higher the abatement of the coalition. As can be seen from Table 4.3 the success of the coalition largely depends on the participation of both USA and CHN (cf. also Dellink 2011).

Table 4.1. Transfers determined by a NBS under different sets of bargaining weights: Grand coalition

Regions	Discount factor	Abatement efforts	Climate change damages	Abatement costs	Economic power
USA	-6364.37	-5687.02	-6209.24	-3275.24	-1312.63
JPN	-7784.00	-12176.82	-11725.37	-12608.12	-8925.65
EUR	-12470.40	-13604.72	-12568.08	-12251.56	-8135.26
OHI	2609.60	-214.57	3603.85	1862.57	1625.68
ROE	1289.63	492.82	2565.22	262.06	1194.35
RUS	3735.38	1109.04	5010.89	2116.59	862.58
HIA	2535.77	1278.78	4356.11	3903.70	1859.10
CHN	2471.66	4876.11	3010.30	4991.20	4383.22
IND	1164.29	1234.09	1584.42	1422.41	583.73
MES	2162.61	2431.62	5254.77	2684.92	1173.04
BRA	4842.53	810.31	7591.57	822.07	1732.48
ROW	1506.56	2181.38	2247.70	3810.50	3079.45
Total	0	0	0	0	0

Note: All figures are expressed as NPV of transfers (bln$) over 100 years

Table 4.2. Stable coalitions and values of bargaining weights under different determinants of bargaining power

Scenario	Stable coalitions	The value of bargaining weights for each region											
		USA	JPN	EUR	OHI	ROE	RUS	HIA	CHN	IND	MES	BRA	ROW
Equal weight	{RUS,CHN, BRA}						0.33		0.33			0.33	
	{HIA,CHN, BRA}							0.33	0.33			0.33	
	{CHN, IND, BRA}								0.33	0.33		0.33	
Discount factor	{ROE, CHN, BRA}					0.34			0.32			0.34	
	{RUS, CHN,BRA }						0.34		0.32			0.34	
Abatement efforts	{OHI, CHN}				0.36				0.64				
	{ROE, CHN}					0.40			0.60				
	{RUS, CHN}						0.43		0.57				
	{HIA, CHN}							0.47	0.53				
	{CHN, IND}								0.47	0.53			
	{CHN, ROW}								0.39				0.61
Climate change damages	{ROE, BRA}					0.40						0.60	
	{HIA, BRA}							0.34				0.66	
	{CHN, BRA}								0.43			0.57	
	{IND, BRA}									0.56		0.44	
	{MES, BRA}										0.42	0.58	
Abatement costs	{OHI, CHN}				0.13				0.87				
	{ROE, CHN}					0.14			0.86				
	{RUS, CHN}						0.20		0.80				
	{CHN, IND}								0.73	0.27			
	{CHN, MES}								0.80		0.20		
	{CHN, ROW}								0.61				0.39
Economic power	{USA, HIA, CHN, BRA}	0.67						0.10	0.16			0.07	
	{CHN, ROW}								0.50				0.50

Table 4.3. Performances of best-performing stable coalitions under different sets of bargaining weights

Regions	Equal weight		Discount factor		Abatement efforts		Climate change damages		Abatement costs		Economic power	
	Annual abatement in 2011 (% of BAU)	NPV of Payoffs (bln$) over 100 years	Annual abatement in 2011 (% of BAU)	NPV of Payoffs (bln$) over 100 years	Annual abatement in 2011 (% of BAU)	NPV of Payoffs (bln$) over 100 years	Annual abatement in 2011 (% of BAU)	NPV of Payoffs (bln$) over 100 years	Annual abatement in 2011 (% of BAU)	NPV of Payoffs (bln$) over 100 years	Annual abatement in 2011 (% of BAU)	NPV of Payoffs (bln$) over 100 years
USA	6.16	4358.32	6.16	4212.30	6.16	4265.77	6.16	3722.80	6.16	4265.77	**7.80**	**4127.30**
JPN	6.31	5284.89	6.31	5120.88	6.31	5180.30	6.31	4556.88	6.31	5180.30	6.31	7078.05
EUR	12.66	6416.46	12.66	6206.67	12.66	6282.83	12.66	5488.42	12.66	6282.83	12.66	8699.96
OHI	1.92	567.84	1.92	549.74	1.92	556.40	1.92	489.74	1.92	556.40	1.92	758.66
ROE	7.93	444.38	7.93	430.24	7.93	435.43	7.93	383.14	7.93	435.43	7.93	594.17
RUS	11.50	1026.68	**14.73**	**885.32**	11.50	1006.30	11.50	884.80	11.50	1006.30	11.50	1376.15
HIA	0.97	568.44	0.97	550.80	0.97	557.21	0.97	490.48	0.97	557.21	**9.35**	**736.23**
CHN	**9.49**	**297.13**	**12.30**	**294.38**	**11.41**	**317.10**	**8.42**	**284.65**	**11.41**	**315.84**	**21.40**	**430.12**
IND	**19.01**	**193.25**	14.85	216.50	14.85	219.40	14.85	192.69	14.85	219.40	14.85	299.96
MES	1.45	413.58	1.45	400.83	1.45	405.45	1.45	356.97	1.45	405.45	1.45	553.02
BRA	**1.75**	**296.22**	**3.01**	**285.34**	0.60	293.52	**1.34**	**254.51**	0.60	293.52	**8.15**	**441.76**
ROW	3.80	1007.11	3.80	974.65	**5.09**	**831.58**	3.80	864.95	**5.90**	**833.85**	3.80	1355.85
Global	7.95	20874.3	8.60	20127.65	8.25	20351.28	7.35	17970.01	8.25	20352.29	11.79	26451.22

Note: Numbers in bold indicate the performances of signatories in a stable coalition for a given set of bargaining weights.

4.4.3 Results for the Nash bargaining solution with outside options

We introduced outside options to the NBS in Section 4.3 and discussed their role. We have argued that the NBS falls into the class of optimal sharing rules when outside options are considered. In this section, we examine the effects of outside options by comparing results from the STACO model for bargaining with outside options with the Nash bargaining outcomes of the previous subsection.

Under the NBS with outside options, signatories' redistributed payoffs consist of their outside option payoff plus a share of the remainder ($\sum_{i\in S} V_i(S) - \sum_{i\in S} V_i(S\backslash\{i\})$). As noted earlier, the distribution of the remainder does not affect coalition stability. Thus stability is independent of the bargaining weights when outside options matter. Our numerical results show a large improvement of the NBS if outside options matter. Both number and size of stable coalitions under the NBS with outside options can be improved as compared to the results for the NBS without outside options. There are more than 190 stable coalitions. Stable coalitions comprise up to six members (e.g. {EUR, OHI, ROE, CHN, MES, BRA}, {EUR, ROE, HIA, CHN, MES, BRA}, {EUR, ROE, HIA, IND, MES, BRA}). This comparison confirms the advantage of the optimal sharing rule in reducing players' free-rider incentives (cf. Weikard and Dellink 2014).

Since outside options are independent of the bargaining weights, any set of bargaining weights will lead to the same set of stable coalitions. Hence there is no need to report results for different sets of bargaining weights in Table 4.4. The best-performing stable coalition is {EUR, CHN, IND, ROW}. Table 4.4 reports the comparison of the two best-performing stable coalitions that are formed under the NBS with and without outside options respectively. There are four members in each. In the case of bargaining with outside options several large GHG emitters (EUR, CHN, IND) are engaged. The high abatement achieved not only generates large net gains for the coalition but also brings significant positive externalities for outsiders. However, due to large free-riding incentives, the best-performing coalition is still of limited size. This enhancement of the abatement efficiency also confirms the numerical results of the optimal sharing scheme obtained by Carraro et al. (2006) and Weikard and Dellink (2014). As shown in the last row of Table 4.4, we use an 'indicator of success' to represent the coalition's efficiency in closing the welfare (or abatement) gap between All Singletons and Grand Coalition[5]. The stable coalition generated by the NBS with outside

[5] The performance of Grand Coalition is reported in Table 4.A3 in Appendix.

options shows a significant improvement in decreasing the gap as compared to the coalition generated without outside options (i.e. 36% and 23%).

Table 4.4 The comparison of two best-performing stable coalitions that are formed under the NBS with and without outside options

Regions	The best-performing stable coalition under NBS with outside options: {EUR, CHN, IND, ROW}		The best-performing coalition without outside options: {USA, HIA, CHN, BRA}	
	Annual abatement in 2011 (% of BAU)	NPV of Payoffs (bln$) over 100 years	Annual abatement in 2011 (% of BAU)	NPV of Payoffs (bln$) over 100 years
USA	6.16	7542.84	**7.80**	**4127.30**
JPN	6.30	8954.01	6.31	7078.05
EUR	**15.33**	**7548.66**	12.66	8699.96
OHI	1.92	958.24	1.92	758.66
ROE	7.93	750.84	7.93	594.17
RUS	11.50	1741.76	11.50	1376.15
HIA	0.97	960.85	**9.35**	**736.23**
CHN	**25.25**	**449.87**	**21.40**	**430.12**
IND	**39.22**	**329.31**	14.85	299.96
MES	1.45	698.92	1.45	553.02
BRA	0.60	505.80	**8.15**	**441.76**
ROW	**21.23**	**1447.48**	3.80	1355.85
Global	15.95	31888.57	11.79	26451.22
Indicator of success (%)	29.38	36.00	15.90	23.00

Note: Numbers in bold indicate the performances of signatories in a stable coalition. Indicator of success (%): (NPV of global payoffs in a coalition – NPV of global payoffs in All Singletons) / (NPV of global payoffs in Grand coalition – NPV of global payoffs in All Singletons) *100. A similar definition applies to abatement.

4.5 Discussions and conclusions

In this paper, we examine the formation and performance of international climate agreements in a cartel game when the distribution of coalitional gains is based on the NBS. We consider different plausible sets of bargaining weights. Our analysis identifies and discusses some key factors driving heterogeneous negotiators' bargaining power in international climate negotiations for distributing cooperative gains. These potential determinants provide insights into countries' potential bargaining positions based on their different characteristics. Furthermore, we consider outside options in the Nash bargaining solution and discuss their role in improving the positive effect of the NBS on the formation and efficiency of ICAs. Our numerical analysis employs the STACO model to investigate the impact of the NBS with asymmetric bargaining power on the formation and efficiency of ICAs.

Firstly, by applying the NBS without outside options to distribute coalitional gains, players' incentives to participate and abate can be increased, although to a limited degree. The effects vary under different sets of bargaining weights. As numerical results in subsection 4.4.2 show, only small coalitions can be stabilised when bargaining weights are determined according to abatement efforts, abatement costs and climate change damages. Our result is in line with the stability results of Weikard et al. (2006), where the coalitional surplus is shared among signatories based on different exogenous claims. In contrast, the size and performance of stable coalitions can be improved when bargaining weights are determined according to the discount factor or economic power.

Secondly, the NBS with outside options is more conducive to ICAs as compared to the bargaining solution without considering outside options. As discussed in subsection 4.3.6, the bargaining outcome falls into the class of optimal sharing rules when outside options are considered. The numerical results in terms of stability and performances of international climate coalitions under the bargaining solution with outside options underline the advantage of such transfers. Our analysis provides a rationale for the use of optimal sharing rules: they result from a NBS with outside options (cf. Muthoo 1999).

Thirdly, multiple equilibrium climate coalitions can emerge from the NBS. In particular under bargaining with outside options we find a large number of equilibrium coalitions. This finding is comparable to Carraro et al. (2006) and Nagashima et al. (2011), where multiple equilibrium coalitions can form when optimal sharing schemes are implemented.

Moreover, it turns out that by applying the NBS to the distribution of coalitional gains, the success of international climate agreements depends on the set of bargaining weights that matters in climate negotiation. Our analysis suggests that some sets of bargaining weights generate more successful coalitions in terms of welfare and abatement than others. For example, among five sets of asymmetric bargaining weights, the one determined by negotiators' economic power can facilitate a climate coalition that comprises two of the largest emitters (CHN and USA) jointly with two other regions.

Our study has some immediate policy implications. Firstly, an ICA should be designed to attract large GHG emitters. Generally, regions with higher GDP produce more emissions, like USA, China or India. Thus when regions with more economic power (higher GDP) can benefit more from an agreement, they will have stronger incentives to join and, hence, more successful ICAs can be formed. It might be controversial that economic power shapes negotiations and determines outcomes. However, it should be noticed that economically powerful regions with a high GDP may include regions with a relatively low GDP per capita, such as China or India when compared to USA. Secondly, in the negotiation process multiple determinants of bargaining power will play a role. This is because one country's incentives to cooperate on GHG mitigation are impacted in a complex way by factors that are related to abatement options, climate change vulnerability and economic power. The bargaining power of each negotiator is likely to be driven by multiple determinants.

One direction to extend our analysis is to study negotiators' strategic behaviour when bargaining power becomes an endogenous variable [6]. This requires an extended dynamic game setting where pre-negotiations determine the negotiation protocol (Wangler et al. 2013) and thereby bargaining powers are relevant at the later stages of the game.

4.6 Appendix

Table 4.A1. Regional aggregation in the STACO3

STACO3	Names
USA	United States
JPN	Japan
EUR	EU27 & EFTA
OHI	Other High Income
ROE	Rest of Europe
RUS	Russia
HIA	High Income Asia
CHN	China
IND	India
MES	Middle East
BRA	Brazil
ROW	Rest of the World

[6] We thank an anonymous reviewer for this comment.

Table 4.A2. Discount rate, GDP and abatement benefit share of twelve regions in STACO 3.0

Regions	Discount rates	GDP in the year 2011(Billion $)	Regional shares of benefits (θ_i)
USA	0.0517	12807.0	0.2263
JPN	0.0359	4831.4	0.1725
EUR	0.0388	13708.0	0.2491
OHI	0.0636	1672.5	0.0345
ROE	0.0612	615.2	0.0271
RUS	0.0397	729.4	0.0403
HIA	0.0474	1973.4	0.0300
CHN	0.1117	3160.0	0.0620
IND	0.1444	803.2	0.0500
MES	0.0470	827.7	0.0249
BRA	0.0442	1266.0	0.0153
ROW	0.0530	3158.6	0.0680
Global	--------	45552.4	$\sum \theta_i = 1$

Table 4.A3. All Singletons and Grand Coalition

Regions	All Singletons		Grand Coalition	
	Annual abatement in 2011 (% of BAU emissions)	Net present value (NPV) of payoffs (Billion $) over 100 years	Annual abatement in 2011 (% of BAU emissions)	Net present value (NPV) of payoffs (Billion $) over 100 years
USA	6.16	3507.55	25.87	12795.98
JPN	6.30	4309.52	23.88	18569.78
EUR	12.66	5173.23	32.13	21508.83
OHI	1.92	463.33	37.08	1429.88
ROE	7.93	362.42	50.64	911.18
RUS	11.50	836.58	50.13	2292.19
HIA	0.97	464.00	29.85	609.19
CHN	6.77	283.53	39.64	-1353.45
IND	14.85	182.14	58.38	21.21
MES	1.45	337.73	73.04	367.20
BRA	0.60	244.52	20.23	393.91
ROW	3.80	816.73	40.94	1027.70
Global	6.88	16981.28	37.73	58573.6

Chapter 5

Compliance and stability of international climate agreements with costly enforcement*

Compliance with abatement commitments is essential for the success of international climate agreements. Due to incentives for free-riding and limited observability of compliance levels, however, an enforcement mechanism including monitoring and potential punishment is required to force members to comply with their committed targets. In this paper we study a sequential cartel formation game, in which the coalition chooses an enforcement policy comprising an abatement target, monitoring expenditures and fines. Individual signatories respond by choosing their preferred abatement level which may or may not comply with the target. In equilibrium, signatories' compliance levels are determined by individual welfare maximization under the agreed enforcement policy. Considering partial compliance, our analysis shows how the extent of compliance depends on abatement targets, monitoring expenditures (the intensity of monitoring) and the fine. Furthermore, we examine the impact of costly enforcement on the stability and performance of international climate agreements. We find that the compliance level of a coalition member can always be improved by increasing the monitoring expenditure. However, the effect of the target on compliance levels depends on the structure of the fine function. Because monitoring is costly, full compliance will usually not be enforced. We also find that a "broad" and "deep" climate coalition cannot be stable under a costly enforcement mechanism.

* This Chapter is based on a manuscript: Yu S., Finus, M., Weikard H.-P.(2016). Compliance and Stability of International climate agreements with costly enforcement. To be submitted.

5.1 Introduction

Incentives to free-ride have been the main obstacle for a successful international cooperation on climate change mitigation. The Paris Agreement adopted in December 2015 seeks to establish broad participation to limit free-riding. However, the success of an international climate agreement (ICA) also depends on the implementation of abatement commitments by signatories. Even after signing a climate treaty, a signatory may have incentives to disregard its abatement obligations. To improve the success of an ICA, signatories need to be incentivised to comply with their commitments.

The idea that contracts need enforcement is not new (e.g. Buchanan 1975). Barrett (2008) has argued that enforcement mechanisms are essential and imperative to ensure the effectiveness of ICAs. When reviewing and evaluating the enforcement mechanism adopted by the Kyoto Protocol, Finus (2008) finds positive effects but also proposes measures to improve enforcement. Based on Finus (2008), Hovi et al. (2012) formulate a pragmatic and credible compliance enforcement system for post-Kyoto climate agreements. Nevertheless, most theoretical and applied analyses of the formation of ICAs only consider free-riding incentives in the participation stage by implicitly assuming models that take compliance for granted (Carraro and Siniscalco 1993; Barrett 1994; Finus and Rundshagen 1998; Weikard et al. 2006; Altamirano-Cabrera and Finus 2006; De Zeeuw 2008; Nagashima et al. 2009). An exception is McEvoy and Stranlund (2009) who study compliance in an ICA formation game that includes a costly monitoring system supported by coalition members. Their theoretical results show that under costly monitoring the set of stable ICAs is smaller but stable coalitions can reach higher levels of participation and abatement compared to costless enforcement where compliance is taken for granted. Based on the same game structure, McEvoy and Stranlund (2010) study the effect of costly enforcement on the efficiency of voluntary environmental agreements, they find that a voluntary environmental agreement can be more efficient in reaching emissions targets than an emissions tax under the condition that the agreement is enforced by a third party that is financially supported by the members of the agreement. Their results also imply that free-riding incentives can be reduced if signatories bear enforcement costs. The analyses of McEvoy and Stranlund (2009, 2010) provide insights into the design of enforcement mechanisms in ICAs where an effective enforcement should be undertaken by an independent third party and funded by all signatories.

However, the analysis of McEvoy and Stranlund (2009, 2010) is simplified as it is based on the assumption that members of an agreement provide sufficient funds to secure full compliance by all members. Abatement decisions are restricted to whether or not to control emissions, resulting in either full compliance or full defection with no control of emissions at all. In other words, their assumption does not consider partial enforcement that may result from signatories' choice of enforcement expenditures. Signatories of ICAs are motivated by cooperative gains to support the enforcement, however they will not contribute an amount that is beyond their benefits from compliance. If their contribution to enforcement is not sufficient for full compliance, partial noncompliance may result. A social optimum can be reached by full compliance of signatories in ICAs. When enforcement is costly, however, the optimal enforcement could induce partial compliance (Arguedas 2008). Stranlund (2007) examines the optimal compliance level in an emissions trading programme. He concludes that under an increasing marginal penalty, some degrees of violation can be cost-effective. In a model where firms can choose cleaner technologies in exchange for reductions of the fines due to non-compliance with environmental standards, Arguedas (2005) finds that the optimal environmental policy could result in a certain degree of non-compliance.

In this paper, we examine an enforcement mechanism similar to the one introduced by McEvoy and Stranlund (2009, 2010) but we consider partial compliance, implying that the optimal abatement level of signatories could lie between the non-cooperative Nash level and the target. With partial compliance, abatement targets might influence the efficiency of ICAs through affecting signatories' compliance levels, and also the degree of sanction on noncompliance. This implies that abatement targets of ICAs should be set optimally in a cooperative way. We relax the assumption of full compliance with the targets of an ICA and determine the optimal enforcement mechanism with respect to the optimal abatement target and monitoring expenditures for an ICA. Furthermore, we analyse how an optimally designed enforcement mechanism affects incentives for participation and compliance in a coalition formation game.

According to the theory of optimal law enforcement (Becker 1968; Polinsky and Shavell 2000), rules are enforced if the expected fine is (weakly) larger than the compliance cost of a potential offender. The expected fine for non-compliance is the probability of detection times the fine. While the sanction is costless in terms of social welfare (if it can be seen as a monetary transfer), detection requires costly monitoring. Therefore, the fine imposed on defectors should be set as high as possible to keep monitoring costs low. The punishment considered in our model is a fine that is dependent on the degree of non-compliance represented by the difference between actual abatement and the negotiated abatement target. To simplify the analysis, we assume that fines collected are paid out to all signatories in a lump-sum way, such that the repayment will not exert influence on signatories' strategic choices of compliance levels. In the model of this paper, the optimal compliance level of signatories is determined endogenously and strategically based on abatement targets, monitoring expenditure and fines imposed on defectors. The optimal level of non-compliance can be explained by the trade-off between gains from increased coalitional abatement due to improved compliance levels and the cost savings from reducing expenditures for monitoring.

Stranlund (2007) and Arguedas (2008) study the effect of the shape of the penalty function on the choice of compliance level. In this paper we study the role of the penalty function for the formation of ICAs and for compliance. In particular, we derive the equilibrium conditions for the case when the fine is linear or convex in the degree of the violation. The monitoring probability incentivises compliance and participation. Hence, we also explore the impact of the monitoring technology by considering monitoring probability as a linear or concave function of monitoring expenditure.

The model in this paper is formulated as a four-stage coalition formation game. At the first stage, countries make their membership choices. At the second stage, coalition members jointly and simultaneously fix their mitigation targets and the monitoring expenditure that determines the inspection probability. At the third stage, countries choose their abatement levels independently. Singletons just choose a best response while signatories must take the expected punishment into account, in case they would not comply. At the final stage, signatories' abatement levels are randomly monitored and fines are due if non-compliance is detected. In the next section we provide details of the game structure and we determine analytically the abatement targets, the optimal monitoring expenditure, individual abatement levels and fines for a given set of signatories.

We find that the compliance level of a coalition member can be improved by increasing the monitoring expenditure. Because monitoring is costly, full compliance will usually not be enforced. If the fine function is linear in the compliance level, it is optimal to set the target at the coalitional optimum. If the fine function is convex, stronger incentives can be set with a higher target. Therefore the target is set at the upper bound which is, in the context of our model, the Business-as-usual (BAU) emissions level. As shown by our theoretical analysis, under the optimal enforcement policy, in equilibrium signatories' abatement level is lower than or equal to the coalitional best level that is obtained from Samuelson's rule. Hence, the optimal enforcement policy induces partial compliance and depends on fine functions and monitoring technologies. Our numerical results show that with a quadratic fine and a linear monitoring probability functions, signatories' can be incentivised to choose the coalitional best abatement level. Otherwise, with constant marginal fine, equilibrium abatement choices will be lower than the coalitional best level. Monitoring technology plays a role for incentivising signatories' compliance level. As our numerical results indicate, signatories' compliance level is lower when the monitoring probability is concave in monitoring expenditures, compared to the case of a linear monitoring probability function.

Fine functions, the enforcement parameters representing the productivity of monitoring expenditure and the severity of the punishment respectively affect the stability of an ICA. The intuition is that stricter enforcement implies higher levels of compliance. This increases costs of compliance due to increased abatement efforts and, in turn, reduces incentives to join an ICA. The converse is also true. Weak enforcement might stabilise a grand coalition. As shown by our numerical results, in the case of quadratic fine and linear monitoring probability functions stable coalitions have the fewest members, while slightly larger coalitions can be stabilised if the fine function is linear or the monitoring probability function is concave.

The rest of the paper is organized as follows. In Section 5.2, we present the model and analyse the equilibrium conditions for each stage of the model. In Section 5.3, under the assumption of symmetric players, we firstly derive the equilibrium for three specified models. Based on the analytical solutions for three models, we illustrate the results on welfare and stability employing a numerical example. We conclude in Section 5.4.

5.2 A general ICA formation model with costly enforcement

The formation of an ICA with costly enforcement is modelled as a four-stage cartel game. Let $N = \{1,2,\dots,n\}$ denote the set of n players. Each player $i \in N$ is faced with a membership choice at stage 1. Signatories, those who sign up to the agreement, form a coalition $S \subseteq N$. At stage 2, all signatories $i \in S$ cooperatively set the abatement targets, denoted by $\bar{q}_i$, for each member in order to maximise joint welfare. Simultaneously, the overall monitoring expenditures denoted by $m = \sum_{i\in S} m_i$ is chosen jointly and optimally by all members through balancing the expenditures of monitoring with benefits from increased compliance. We assume in our model that the aggregated monitoring expenditures m are shared equally among coalition members, i.e. $m_i = \frac{1}{s}m$, hence we focus on how to determine the optimal aggregated level of monitoring expenditure m. At stage 3, with given abatement targets and the monitoring expenditure, coalition members and singletons choose abatement levels q_i by maximising their individual welfare. The equilibrium solution determines the level of compliance. At the final stage, each signatory's abatement is randomly monitored with a probability $P \in [0,1]$ depending on the monitoring expenditure m and the size of the coalition $s = |S|$. Thus the inspection probability can be represented as a function of the monitoring "intensity" $\frac{m}{s}$ such that $P = P(\frac{m}{s})$ with $P(0) = 0$, $\frac{\partial P}{\partial m} > 0$ and $\frac{\partial P}{\partial s} < 0$. If emission reduction by any signatory is found to be less than its abatement commitment, a fine will be imposed.

We apply sub-game perfect equilibrium to solve the game, such that equilibria are obtained by backward induction.

Stage 4: Starting with the analysis of the final stage, notice that at this stage the set of signatories S, the monitoring expenditure m, and the signatories abatement q_i and abatement targets $\bar{q}_i$ are given. All signatories are monitored by an enforcement agency with probability $P(\frac{m}{s})$. A fine, denoted by F_i, is imposed on the defector if non-compliance is detected. The fine is increasing in the level of defection, i.e. the shortfall of a country's abatement compared to its target. Let $d_i \equiv max(0, \bar{q}_i - q_i)$ denote the level of defection. The max operator ensures that overcompliance does not count as "negative defection". We also introduce a parameter $f > 0$ which reflects the severity of punishment. Here we assume that the fine imposed would reflect the severity of the offence; it could, for example, reflect the (global) damage from emissions. Note that, as Stranlund (2007) points out, different forms of penalty influence the choice of abatement in the way that the abatement depends on the punishment policy parameter f or abatement target $\bar{q}_i$ or both. In our model we treat f as an exogenous parameter that is set in pre-negotiations. Then a general way to write the fine is $F_i(d_i; f) = F_i(\bar{q}_i, q_i; f)$. Fines are assumed to be weakly convex in the level of defection; there is no discount on punishment for more severe defections. Given the choices in previous stages, signatories' expected fine is $\hat{F}_i = P(\frac{m}{s}) \cdot F_i(\bar{q}_i, q_i; f)$.

Stage 3: At stage 3, countries choose their abatement levels strategically. Any player's abatement cannot exceed the Business-as-usual (BAU) emissions level denoted by $\bar{e}_i$, such that $q_i \in [0, \bar{e}_i]$. Let B_i and C_i denote respectively abatement benefit and cost functions where the magnitude of abatement benefits depends on the global abatement denoted by $q = \sum_{i \in N} q_i$.

We can write $B_i = B_i(q)$ with $B_i'(q) > 0$ and $B_i''(q) \leq 0$. The cost is associated with individual abatement $C_i(q_i)$ with $C'_i(q_i) > 0$ and $C''_i(q_i) > 0$. At this stage, singleton players choose their abatement level to maximise their own payoffs by taking all others' abatement as given. Under such an abatement denoted by q_i^{NS}, in equilibrium each non-signatory's marginal abatement benefits are equal to its marginal costs:

$$B_i'(q) = C_i'(q_i^{NS}), i \in N \backslash S. \tag{5.1}$$

Each signatory decides on its abatement level to maximize its own expected payoff given the coalition's monitoring expenditure m and abatement targets $\bar{q}_i$ decided at stage 2. The problem of signatory $i \in S$ at this stage can be written as:

$$\max_{q_i} \pi_i(\boldsymbol{q}) = B_i(q) - C_i(q_i) - m_i - P(\tfrac{m}{s})F_i(\bar{q}_i, q_i). \tag{5.2}$$

where $\boldsymbol{q} = (q_1, \dots, q_n)$, denotes the abatement vector. Notice that an abatement target that is smaller than the Nash abatement level (in the absence of a coalition) denoted by q_i^N would not be effective. Hence, even before we analyse the second stage we can assume targets $\bar{q}_i \geq q_i^N$. The enforcement mechanism cannot incentivise abatement levels beyond $\bar{q}_i$, hence, we can rule out overcompliance such that signatories' abatement $q_i \in [0, \bar{q}_i]$. Note that we assume that fines, if collected, are distributed as a lump sum to all signatories. This assumption implies that the redistributed fine will not influence signatories' strategic choice of abatement. Taking the derivative of problem (5.2) with respect to q_i, the first order condition for an interior solution is obtained as:

$$B_i'(q^*) - C_i'(q_i^*) - P(\tfrac{m}{s}) \frac{\partial F_i(\bar{q}_i, q_i)}{\partial q_i}\bigg|_{q_i = q_i^*} = 0 \Leftrightarrow B_i'(q^*) - C_i'(q_i^*) = P(\tfrac{m}{s}) \frac{\partial F_i(\bar{q}_i, q_i)}{\partial q_i}\bigg|_{q_i = q_i^*}. \tag{5.3}$$

Eq. (5.3) shows how signatories' abatement choice is influenced by the enforcement mechanism. Note that $\frac{\partial F_i(\bar{q}_i, q_i)}{\partial q_i}\Big|_{q_i = q_i^*} \leq 0$, thus the equilibrium abatement of a signatory $i \in S$ is not less than the non-cooperative Nash equilibrium level, where $B_i'(q^*) = C_i'(q_i^*)$. For an interior solution, according to the equilibrium condition (5.3), a signatory chooses the optimal level of abatement by equalising marginal net gains of abatement with the marginal expected fine. Notice that, if the marginal expected fine is zero, no signatory will comply with the target and the non-cooperative abatement level will result. It is also interesting to note that Samuelson's rule which determines the coalitional best level of abatement can be satisfied if

$-P(\frac{m}{s})\frac{\partial F_i(\bar{q}_i,q_i)}{\partial q_i}\Big|_{q_i=q_i^*} = \sum_{j\in S\setminus\{i\}} B_j'$. The intuition is as follows: if the marginal expected fine makes up for the externality of each signatory's abatement, then the coalitional optimum will be achieved.

Since Eq. (5.3) is a condition for an interior solution $q_i^* \leq \bar{q}_i$, it is also possible that

$$B_i'(\bar{q}_i) - C_i'(\bar{q}_i) > P\left(\frac{m}{s}\right)\frac{\partial F_i(\bar{q}_i, q_i)}{\partial q_i}\Big|_{q_i=\bar{q}_i}. \tag{5.4}$$

In this case we have a corner solution, where $q_i^* = \bar{q}_i$ and full compliance is achieved. This happens when the punishment for a small deviation from the target is sufficiently severe.

From Eq. (5.3) we can see that the higher the monitoring probability P, the higher is the abatement q_i^* (and the compliance level). This indicates an implicit relationship between equilibrium abatement of signatories and monitoring probability, that is:

$$\frac{\partial q_i^*}{\partial P} > 0, i \in S. \tag{5.5}$$

We can now write the optimal abatement as the reduced form $q_i^* = q_i^*(\bar{q}_i, P(\frac{m}{s}))$. Since the detection probability P is a function of m with $\frac{\partial P}{\partial m} > 0$, we can write:

$$q_i^*\left(\bar{q}_i, P(\tfrac{m}{s})\right) = q_i^*(\bar{q}_i, m, s), with\ \frac{\partial q_i^*}{\partial m} > 0, i \in S. \tag{5.6}$$

Hence, by raising the monitoring expenditure m, the inspection probability P and the equilibrium abatement q_i^* increase. However, the relationship between q_i^* and $\bar{q}_i$ is not straightforward, as it depends on the form of the fine function and will be discussed in the next section.

Stage 2: Now we move to the second stage. At this stage coalition members jointly determine abatement targets $\bar{q}_i$ and monitoring payment m. Because monitoring is costly and increasing the target is costless, it is always better to increase the target and to lower monitoring efforts while maintaining the expected fine. Unless there is an upper bound of the target, we cannot obtain a solution. However it is reasonable to assume that the target is bounded by the BAU emissions $\bar{e}_i$, i.e. $\bar{q}_i \leq \bar{e}_i$. Note that coalitional payoffs are decreasing in monitoring expenditures. Since we assume fines collected from defectors will be paid back to signatories, the collected fines are welfare neutral from the perspective of the coalition. The coalition solves the following problem:

$$\max_{m,\bar{q}_i} \sum_{j\in S} \pi_j(\boldsymbol{q}^*) = \sum_{j\in S} B_j(q^*) - \sum_{j\in S} C_i\left(q_j^*(\bar{q}_i, m)\right) - m \,. \tag{5.7}$$

s.t.

$\bar{q}_i \leq \bar{e}_i$.

The Lagrangian function for the problem (5.7) is $L(m;\bar{q}_i;\gamma_i) = \sum_{j\in S}\pi_j(\boldsymbol{q}^*) + \gamma_i(\bar{e}_i - \bar{q}_i)$, where γ_i is the Lagrangian multiplier for the target. By taking the partial derivatives with respect to m, $\bar{q}_i$ and γ_i, we obtain the following first order conditions:

$$\frac{\partial L}{\partial m} = \sum_{j\in S}\frac{\partial B_j}{\partial q^*}\sum_{j\in S}\frac{\partial q_j^*}{\partial m} - \sum_{j\in S}(\frac{\partial C_j}{\partial q_j^*}\frac{\partial q_j^*}{\partial m}) - 1 = 0, \tag{5.7a}$$

$$\frac{\partial L}{\partial \bar{q}_i} = (\sum_{j\in S}\frac{\partial B_j}{\partial q^*})\frac{\partial q_i^*}{\partial \bar{q}_i} - \frac{\partial C_i}{\partial q_i^*}\frac{\partial q_i^*}{\partial \bar{q}_i} - \gamma_i = 0, \tag{5.7b}$$

$$\frac{\partial L}{\partial \gamma_i} = \bar{e}_i - \bar{q}_i \geq 0, \tag{5.7c}$$

$$\gamma_i(\bar{e}_i - \bar{q}_i) = 0, \qquad \gamma_i \geq 0. \tag{5.7d}$$

It is clear from (5.7d) that if the constraint is non-binding, then $\gamma_i = 0$ and Samuelson's rule applies according to (5.7b). However, we prove in the Appendix that the optimal target is always chosen as a corner solution, i.e. $\bar{q}_i^* = \bar{e}_i$. Then the first order conditions (5.7a-d) can be reduced to the following:

$$\sum_{j\in S}\frac{\partial B_j}{\partial q^*}\left(\sum_{j\in S}\frac{\partial q_j^*}{\partial m}\right) - \sum_{j\in S}(\frac{\partial C_j}{\partial q_j^*}\frac{\partial q_j^*}{\partial m}) - 1 = 0. \tag{5.8}$$

$$(\sum_{j\in S}\frac{\partial B_j}{\partial q^*} - \frac{\partial C_i}{\partial q_i^*})\frac{\partial q_i^*}{\partial \bar{q}_i} = \gamma_i \geq 0. \tag{5.9}$$

In Eq. (5.8) the term $\frac{\partial q_j^*}{\partial m}$ gives the marginal incentive to abate. Since we assume that overcompliance does not pay, additional monitoring will not increase abatement beyond the target level. Eq. (5.8) indicates that in equilibrium, the coalitional marginal net gains of the increased abatement due an increase in monitoring efforts is equal to the marginal costs of monitoring which are unity by assumption.

Eq. (5.9) leads to interesting insights. First, if $\gamma_i > 0$, then $\sum_{j \in S} \frac{\partial B_j}{\partial q^*} - \frac{\partial C_i}{\partial q_i^*}$ and $\frac{\partial q_i^*}{\partial \bar{q}_i}$ must both be positive. This is true because it is never optimal for a coalition to induce abatement higher than the coalitional best level and we can rule out that $\sum_{j \in S} \frac{\partial B_j}{\partial q^*} - \frac{\partial C_i}{\partial q_i^*} < 0$. Hence, $\gamma_i > 0$ implies that the level of abatement that is optimal to enforce falls short of the coalitional best (Samuelson) level of abatement. Furthermore, from Eq. (5.9) the coalitional best abatement can be induced by the optimal target if $\sum_{j \in S} \frac{\partial B_j}{\partial q^*} - \frac{\partial C_i}{\partial q_i^*} = 0$. The shadow value γ_i indicates the marginal gain if the constraint on the target could be relaxed. Second, note that, from Eq. (5.9) it is also possible that increasing the target is not effective for inducing higher abatement. In that case $\frac{\partial q_i^*}{\partial \bar{q}_i} = 0$. This is the case when the fine function is linear.

From above analysis of Eq. (5.9), it can be concluded that, if the fine function is convex and therefore $\frac{\partial q_i^*}{\partial \bar{q}_i} > 0$, and under the binding target $\bar{q}_i^* = \bar{e}_i$, signatories are incentivised to choose an abatement level that is equal to or lower than the coalitional first best level obtained from the Samuelson's rule, i.e. $\sum_{j \in S} \frac{\partial B_j}{\partial q^*} \geq \frac{\partial C_i}{\partial q_i^*}$. Particularly, the target is set optimally so that signatories could be induced to choose the coalitional best abatement level. As argued before this will be achieved if

$$-P\left(\frac{m}{s}\right) \frac{\partial F_i(\bar{q}_i, q_i)}{\partial q_i}\bigg|_{q_i = q_i^*} = \sum_{j \in S \setminus \{i\}} B_j'. \qquad (5.10)$$

We can also see from (5.10), that the target cannot play a role in incentivising signatories' abatement when the marginal fine is constant. Some implications can be gained from equilibrium condition (5.9). Firstly, under the optimal enforcement mechanism, the target and monitoring are used as two instruments to increase signatories' compliance level. Since setting higher target is costless while increasing the monitoring probability is not, it is always better to choose the maximum target level. Secondly, setting higher targets can increase the abatement, but it is not optimal for a coalition to abate more than the level obtained from the Samuelson condition. Under the Samuelson condition each coalition member makes the abatement choice that internalises the externality imposed on all other signatories.

Stage 1: At the initial stage all countries make decisions on the membership by evaluating their payoffs of being a signatory or a singleton. Countries evaluate payoffs depending on the anticipated strategic decisions on the optimal enforcement policy and the abatement. We assume the sub-game perfect equilibrium in our model is unique under the optimal enforcement policy. Therefore, for each coalition structure we introduce a valuation function $V_i(S)$ to represent each player's payoff under a coalition S. By applying the solution concept of cartel stability (d'Aspremont et al., 1983) to represent the Nash equilibrium at this stage, a stable coalition is defined as:

(a) internal stability: $V_i(S) \geq V_i(S\backslash\{i\}), \qquad \forall i \in S$ (5.11)

(b) external stability: $V_i(S) \geq V_i(S \cup \{i\}). \qquad \forall i \in N\backslash S$ (5.12)

5.3 Models with specified functional forms

However, outcomes and policy implications can be different under different functional forms. For example, the form of fine function F can affect the choice of monitoring expenditures, abatement choices and therefore also signatories' incentives to participate in an ICA. Furthermore, the monitoring technology affects the expected penalty and thus compliance levels. To explore the impact of design features (i.e. the penalty structure and monitoring technology) of enforcement mechanisms in ICAs, this section provides equilibrium conditions obtained from Section 5.2 for models with specified functions. We present three specified models that differ in the form of fine (penalty) and monitoring probability functions. For simplicity, we assume symmetric players with same abatement cost and benefit functions in all specified models. To illustrate more explicitly the impact of the costly enforcement on welfare and stability under these three model specifications, a numerical example is provided in the final part of this section.

Under all three model specifications, the benefit and cost functions remain the same. We employ a linear abatement benefit function. In this case non-signatories have dominant abatement strategies which coincide with the non-cooperative Nash equilibrium. This allows us in the following subsections to focus our analysis on signatories' choices.

5.3.1 A basic model with a linear fine and a linear monitoring probability functions

Our analysis starts with a basic model, where the monitoring probability and the fine are linear. Functions are specified as follows. We assume linear abatement benefits $B_i(q) = bq$

with $b > 0$; marginally increasing abatement cost $C_i(q_i) = \frac{1}{2}aq_i^2$ with $a > 0$; and a linear monitoring probability $P(\frac{m}{s}) = \alpha\frac{m}{s}$ with $0 \leq P \leq 1$. In this function α is a parameter representing the monitoring productivity and satisfying $\alpha > 0$. We assume that all signatories can be monitored with the probability one if the monitoring expenditure is large enough, namely if $m \geq \frac{s}{\alpha}$. The fine imposed on a defector is linear in the degree of non-compliance $F_i = f \cdot (\bar{q}_i - q_i)$ where $f > 0$ is a parameter. Moreover, observe that if $q_i = 0$ and for given $\bar{q}_i$, the maximum possible expected fine is $\alpha\frac{m}{s}f\bar{q}_i$.

According to the equilibrium condition of signatories' abatement choice (5.3) at stage 3, the optimal individual abatement q_i^* of a coalition member $i \in S$ is

$$q_i^* = \frac{b}{a} + \frac{\alpha m f}{as}. \tag{5.13}$$

It is obvious from Eq. (5.13) that for a given coalition S the choice of optimal abatement q_i^* for signatories $i \in S$ only depends on monitoring expenditures m and exogenous parameters. Clearly, the abatement of each signatory is higher than the non-cooperative Nash level $\frac{b}{a}$ as long as the monitoring expenditure is positive. According to Eq. (5.13), for interior solutions $q_i^* < \bar{q}_i$, the target $\bar{q}_i$ cannot affect signatories' abatement choice due to the constant rate of the marginal fine $\left|\frac{\partial F_i(\bar{q}_i, q_i)}{\partial q_i}\right| = f$. However, this conclusion does hold for a corner solution $q_i^* = \bar{q}_i$, where signatories' optimal abatement q_i^* can be advanced by setting higher target $\bar{q}_i$.

As shown by the analysis of the stage 2 game in Section 5.2, a linear fine function implies that the target does not play a role in inducing higher abatement levels that are interior solutions $q_i^* < \bar{q}_i$. However, for a corner solution $q_i^* = \bar{q}_i$, signatories' abatement can be advanced by increasing the target $\bar{q}_i$. In such case optimal abatement can be implemented by setting the target equal to the coalitional best level, i.e. $\bar{q}_i^* = \frac{sb}{a}$[7].

The optimal monitoring expenditures m^* can be determined by applying $\frac{\partial q_i^*}{\partial m} = \frac{\alpha f}{as}$ (which is obtained from Eq. (5.13)), to the equilibrium condition (5.8). The following result can be derived based on specified functions in this basic model:

[7] This value is obtained by applying the specified functions in this basic model to the Samuelson condition $\frac{\partial C_i}{\partial q_i^*} = \sum_{j \in S} \frac{\partial B_j}{\partial q^*}$.

$$q_i^* = \frac{sb}{a} - \frac{1}{\alpha f}. \tag{5.14}$$

Since the term m cancels out, Eq. (5.14) gives signatories' optimal abatement level that is resulted from the optimal monitoring costs. As shown by Eq. (5.14), for given parameters (i.e. α, f, a and b), the coalitional abatement increases with coalition size s. This reflect the well-known result that incentives to become a free-rider are increasing in the coalition size. Moreover, according to Eq. (5.14), the coalition best abatement level $\frac{sb}{a}$ cannot be achieved by an interior solution q_i^* considering $\frac{1}{\alpha f} > 0$. This result can be interpreted as the optimal monitoring costs is not sufficient to induce signatories to choose the coalitional best abatement. As a consequence partial compliance results. Optimal monitoring costs m^*, are obtained by equalizing Eq. (5.14) with Eq. (5.13) and rearranging the equation:

$$m^* = \frac{s(s-1)b}{\alpha f} - \frac{as}{\alpha^2 f^2}. \tag{5.15}$$

From Eq. (5.15) we see that the individual monitoring contribution $m_i^* = \frac{(s-1)b}{\alpha f} - \frac{a}{\alpha^2 f^2}$ increases in the size of a coalition. This, jointly with (5.14) shows that for each individual signatory, compliance costs increase with the size of the coalition. Hence, incentives to participate and to comply become smaller when the coalition is getting larger.

Based on the above analysis on optimal solutions for signatories' and singletons' abatement, monitoring expenditure, the equilibrium payoff of each signatory under the coalition S can be obtained as:

$$V_i^S(s) = \frac{(n-s)b^2}{a} + \frac{s^2 b^2}{2a} - \frac{(s-1)b}{\alpha f} + \frac{a}{2\alpha^2 f^2}, i \in S. \tag{5.16}$$

The singleton's equilibrium payoff is solved as:

$$V_i^{NS}(s) = \frac{(s^2 - s + n)b^2}{a} - \frac{b^2}{2a} - \frac{sb}{\alpha f}, i \in N \backslash S. \tag{5.17}$$

By applying the solution concept of internal and external stability, the formation of ICAs in the first stage can be analysed. We use a stability function (Carraro and Siniscalco 1993) to identify the size of the stable coalition. Generally the stability function is defined as $\Lambda(s) = V_i^S(s) - V_i^{NS}(s-1)$. The size of the stable coalition is the largest integer s satisfying $\Lambda(s) \geq$

0 and $\frac{\partial \Lambda(s)}{\partial s} < 0$. Based on the welfare equations (5.16) and (5.17), the stability function can be specified as:

$$\Lambda(s) = \frac{(4s - s^2 - 3)b^2}{2a} + \frac{a}{2\alpha^2 f^2}. \tag{5.18}$$

Thus the derivative of $\Lambda(s)$ can be obtained as:

$$\frac{\partial \Lambda(s)}{\partial s} = \frac{(2 - s)b^2}{a}. \tag{5.19}$$

Eq. (5.19) shows that the stability function is falling for $s > 2$. This implies that minimum value of the size of the internally and externally stable coalition is 2. From (5.18) we conclude that the coalition size is decreasing in α and f. The grand coalition (i.e. $s = n$) even can be stabilised if the term αf is low enough. The intuition behind this conclusion is that α and f are parameters representing the monitoring productivity and the severity of punishment respectively, higher value of α and f implies a defector faces higher probability of being monitored and having a heavier fine punishment. This increases the effectiveness of an agreement and therefore the cost of defection. Accordingly, a potential participating country will have less incentives to join an ICA in order to avoid strict enforcement policies. Conversely, if enforcement is weak, the grand coalition is stable but abatement is close to the non-cooperative Nash levels. Our conclusion on the impact of the costly enforcement on the membership of an ICA is in line with McEvoy and Stranlund (2009) who also highlight the adverse impact of the costly enforcement.

Taken together, the result of the linear fine model shows that signatories can be induced to increase their compliance levels by increasing monitoring expenditures. Due to the constant marginal fine the target is an ineffective enforcement instrument for interior solutions for signatory's equilibrium abatement, while it can only advance each signatory's abatement level at corner solutions. In equilibrium the monitoring expenditure chosen by the coalition can induce the signatories to choose abatement levels that are lower than the coalitional best levels determined by Samuelson's rule. Consequently, only partial compliance can be sustained by the optimal enforcement policy. As to the impact of the costly enforcement on the membership choice, the size of the stable ICA with costly enforcement is ambiguous and dependent on the value of enforcement parameters. This implies that the grand coalition might exist if the enforcement parameters are properly taken at a value low enough. We resort to a numerical example in Section 5.3.4 to illustrate welfare effects of the costly enforcement mechanism with linear fine and monitoring probability functions.

5.3.2 The model with a quadratic fine function

The structure of the fine is one of the key factors to the endogenous choice of enforcement policies. Stranlund (2007) and Arguedas (2005) conclude that the choice of the optimal compliance level depends on the shape of penalty. In the previous section with the linear fine function, due to the constant expected marginal fines, signatories' choices of compliance levels are independent of the abatement target but only depends on the monitoring expenditures. To investigate the effect of different fine functions on the choice of enforcement policy, we extend the linear fine function to the quadratic form in this section.

By applying $F_i = \frac{1}{2} f(\bar{q}_i - q_i)^2$ and other specified functions that remain the same as in the basic model, the equilibrium condition (5.3) for signatories' abatement can be rewritten as:

$$q_i^* = \frac{bs + \alpha m f \bar{q}_i}{as + \alpha m f}. \tag{5.20}$$

Clearly, Eq. (5.20) shows signatories' abatement can be increased by setting a higher target $\bar{q}_i$. As to the effect of monitoring efforts, because $\frac{\partial q_i^*}{\partial m} = \frac{s\alpha f(a\bar{q}_i - b)}{(as+\alpha m f)^2} > 0$ according to Eq. (5.20) and the assumption $\bar{q}_i \geq q_i^N = \frac{b}{a}$ signatories' abatement level will increase with larger monitoring efforts. It is worth noting that the increasing rate of signatories' abatement driven by monitoring efforts is diminishing due to $\frac{\partial^2 q_i^*}{\partial m^2} = -\frac{2s(s-1)b}{\alpha f m^3} < 0$. It can be concluded from Eq. (5.20) that under the marginally increasing fine, signatories' compliance levels can be advanced by either setting higher targets or spending more on monitoring. As shown in the analysis in Section 5.2, the optimal target is set at its upper bound $\bar{q}_i^* = \bar{e}_i$. Because players are symmetric in our specification, we remove the subscript of $\bar{e}_i$ and then $\bar{q}_i^* = \bar{e}$. By replacing $\bar{q}_i$ in Eq. (5.20) with $\bar{e}$, the optimal abatement level of each signatory can be rewritten as $q_i^* = \frac{bs+\alpha m f \bar{e}}{as+\alpha m f}$.

According to Eq. (5.9), under the optimal target $\bar{q}_i^* = \bar{e}$, signatories are incentivised to choose an abatement that is lower than or equal to the coalitional best level. Hence, signatories' optimal abatement under the quadratic fine structure can be represented as the following:

$$q_i^* = \begin{cases} \dfrac{sb}{a} & if \ \dfrac{bs+\alpha mf\bar{e}}{as+\alpha mf} \geq \dfrac{sb}{a} \\ \dfrac{bs+\alpha mf\bar{e}}{as+\alpha mf} & if \ \dfrac{bs+\alpha mf\bar{e}}{as+\alpha mf} < \dfrac{sb}{a} \end{cases}. \tag{5.21}$$

Note that for both values of q_i^*, signatories only partially comply with the target $\bar{q}_i^* = \bar{e}$.

Now we come to the choice of monitoring expenditures, which are derived differently depending on q_i^* as shown in Eq. (5.21). When the coalitional best abatement is induced $q_i^* = \frac{sb}{a}$, the monitoring costs that are sufficient to complement the maximum target for implementation of the coalitional best abatement level can be directly identified by solving for m in Eq. (5.10). Otherwise under a target that has a positive shadow value, monitoring costs that are coupled with the target to induce higher abatement level can be obtained according to Eq. (5.8) that is the equilibrium condition for choosing the optimal monitoring costs. Hence, by applying specified functions specification to Eqs. (5.8) and (5.10), the optimal monitoring costs m^*can be solved as:

$$m^* = \begin{cases} \dfrac{s(s-1)ab}{\alpha f(a\bar{e}-sb)} & if \ q_i^* = \dfrac{sb}{a} \\ \dfrac{s\left(\sqrt{\dfrac{\alpha f(a\bar{e}-b)(abs^2+\alpha m^* fsb - abs - \alpha m^* fa\bar{e})}{as+\alpha m^* f}} - a\right)}{\alpha f} & if q_i^* = \dfrac{bs+\alpha m^* f\bar{e}}{as+\alpha m^* f} \end{cases}. \tag{5.22}$$

It can be seen from Eq. (5.22) that larger monitoring expenditures will be required for a low monitoring technology α under the coalitional best abatement level. However the optimal monitoring expenditures under the abatement $q_i^* < \frac{sb}{a}$ cannot be derived explicitly, and we resort to a numerical example in Section 5.3.4.

As shown by the above analysis, for the case of $m^* = \frac{s(s-1)ab}{\alpha f(a\bar{e}-sb)}$ and $q_i^* = \frac{sb}{a}$, we can derive the stability function as the following. The payoff of each signatory is obtained as:

$$V_i^S(s) = \frac{s^2b^2}{2a} - \frac{sb^2}{a} + \frac{nb^2}{a} - \frac{(s-1)ab}{\alpha f(a\bar{e}-sb)}, i \in S. \tag{5.23}$$

The singleton's optimal payoff is solved as:

$$V_i^{NS}(s) = \frac{s^2 b^2}{a} - \frac{sb^2}{a} + \frac{(2n-1)b^2}{2a}, i \in N \backslash S. \tag{5.24}$$

Based on these welfare function, the stability function is:

$$\Lambda(s) = V_i^S(s) - V_i^{NS}(s-1) = \frac{(-s^2+4s-3)b^2}{2a} - \frac{(s-1)ab}{\alpha f(a\bar{e} - sb)}. \tag{5.25}$$

The size s satisfying conditions $\Lambda(s) \geq 0$ and $\frac{\partial \Lambda(s)}{\partial s} < 0$ cannot be solved explicitly, hence we have to resort to a numerical example in Section 5.3.4 to show the stability effect.

As compared to the result of the basic model, due to increased severity of punishment, signatories can be incentivised to change their compliance levels by both of monitoring costs and the fine through changing of the target. Under the optimal target chosen at the upper bounder, signatories only partially comply with the target by choosing the abatement that is lower than or equal to the coalitional best level. The effect of the enforcement with marginally increasing fine function on the stability and welfare of an ICA will be checked by resorting to a numerical example in Section 5.3.4.

5.3.3 The model with a concave function of monitoring probability

Under the linear monitoring probability function, the monitoring probability for each signatory is increasing constantly with the monitoring expenditure. According to the law of diminishing returns, however, the marginal productivity of the monitoring expenditure referring to the marginal monitoring probability, could be decreasing with monitoring expenditures. In this section, we investigate the enforcement policy with a concave form of probability function while keeping other functions unchanged as compared to the basic model. Considering the maximum monitoring probability and the diminishing returns of the monitoring expenditure $\frac{m}{s}$, the function of monitoring probability is specified as $P\left(\frac{m}{s}\right) = 1 - \frac{1}{1+\alpha\frac{m}{s}}$. By applying specified functions to the equilibrium condition (5.3) at stage 3, the strategy for the optimal abatement choice of each signatory is derived as:

$$q_i^* = \frac{b}{a} + \frac{\alpha m f}{a(\alpha m + s)}. \tag{5.26}$$

Because the fine function is linear in this case, as shown in Eq. (5.26) signatories' abatement level can only be incentivised by the monitoring expenditure. Furthermore, according to Eq.

(5.26) it can be shown that signatories' abatement is increasing at a decreasing rate, i.e. $\frac{\partial q_i^*}{\partial m} = \frac{\alpha f s}{a(\alpha m+s)^2} > 0$, $\frac{\partial^2 q_i^*}{\partial m^2} = -\frac{2\alpha^2 f s}{a(\alpha m+s)^3} < 0$. It is also obvious to see from Eq. (5.26) that signatories' optimal abatement level is higher than the Nash level $\frac{b}{a}$ given the term $\frac{\alpha m f}{a(\alpha m+s)} > 0$ under positive monitoring expenditures, but all signatories will become non-compliant and their abatement will equal to the Nash level when there is no monitoring expenditure $m = 0$. By equalizing the coalitional best abatement level $\frac{sb}{a}$ with q_i^* shown in Eq. (5.26) and then solving for m, we can see that the coalition will reach the fist-best if $m = \frac{s(s-1)b}{\alpha(f+b-sb)}$, which can serve as an upper bound for the monitoring expenditure, i.e. $m \leq \frac{s(s-1)b}{\alpha(f+b-sb)}$. This is because it is not coalitional optimal to spend more on monitoring if the coalitional best abatement is taken by signatories, as shown by the equilibrium condition (5.9). But note that, this upper bound level that can induce the coalitional best abatement is not the optimal level of the monitoring expenditure in a coalition.

Following the target setting under the model with linear fine function in Section 5.3.1, the optimal target is set as $\bar{q}_i^* = \frac{sb}{a}$. Applying specified functions to Eq. (5.8) that is the condition for the optimal monitoring expenditure, and subtracting m yield the following result:

$$m^* = \frac{s}{\alpha}\sqrt{\frac{b\alpha f(s-1) - \frac{m^*\alpha^2 f^2}{\alpha m^* + s}}{a}} - \frac{s}{\alpha}\,. \tag{5.27}$$

Eq. (5.27) is a cubic function of m^*, implying an explicit analytical solution for optimal monitoring expenditure m^*cannot be obtained. We resort to an numerical example in Section 5.3.4 to show the optimal monitoring expenditure, and also results on the compliance level and the stability.

5.3.4 A numerical illustration

Implicit solutions for the optimal monitoring cost m^* in Sections 5.3.2 and 5.3.3 cannot enable us to look into the effect of the optimal enforcement mechanism in an ICA. In this section we use a numerical example to show the optimal monitoring cost choice and the impact of which on signatories' compliance levels (abatement levels), the coalition stability and payoffs. We set the following parameter values throughout this numerical example: the number of players $n = 10$, the BAU emissions $\bar{e} = 110$, the abatement cost parameter $a = 1$, the marginal abatement benefit $b = 10$, the parameter representing the severity of the penalty in the fine function $f = 100$, the parameter representing monitoring technology in the monitoring probability function $\alpha = 0.01$. Numerical results are shown in the following tables.

Let us firstly look into the result on the compliance level under each model specification. As shown in row 7 of Table 5.1, under the enforcement mechanism with linear fine function, signatories' optimal abatement levels q_i^* under different coalition sizes are below but close to the coalitional best level (i.e. $\frac{sb}{a}$). Under fines that are marginally increasing in defection levels, the results of signatories' optimal abatement in Table 5.2a show that the coalitional best abatement level $q_i^* = \frac{sb}{a}$ can be induced by the optimal enforcement policy composing both of the optimal target and monitoring expenditures. For the case $q_i^* = \frac{bs+\alpha mf\bar{e}}{as+\alpha mf} < \frac{sb}{a}$, each signatory's abatement level is slightly lower than the first-best level. However, the differences in the optimal abatement levels and monitoring expenditures between these two cases are becoming larger when the coalition is expanding, i.e. after $s = 7$. The changes in these two cases can be attributed to the functional form of the analytical solutions (i.e. Eqs. (5.21) and (5.22) representing m^* and q_i^*), and also because of the particular numerical values given in this example. As shown in row 7 of Table 5.2a and 5.2b, in general the optimal monitoring expenditures under the case $q_i^* = \frac{sb}{a}$ are also slightly higher than the case $q_i^* < \frac{sb}{a}$. This is because as analysed in Section 5.3.2, under quadratic fine function monitoring costs are chosen at a value which can work together with the optimal target to incentivise signatories to choose the coalitional best abatement level. For the case where the optimal abatement level is lower than the coalitional best level, the monitoring costs are also lower. When the monitoring technology parameter is low, as shown in Table 5.3 where the marginal monitoring probability is decreasing and the marginal fine is constant, in equilibrium the marginal abatement level that is induced by one more unit of monitoring expenditure is the lowest as compared to the other two models with constant marginal monitoring probability. Because of the low efficiency of the enforcement with concave monitoring probability, monitoring costs are more than half of signatories' payoffs when the coalition size is larger than eight. The optimal targets under these three model specifications are all set at the BAU emissions level, hence signatories are partially complying with the target under these three cases. As shown by the numerical results, the compliance level and payoffs under the quadratic fine and constant marginal monitoring probability are the highest while the ones under the linear fine and the concave monitoring probability functions are the lowest.

For the results on coalition stability under three model specifications, it can be seen from Tables 5.1-5.3 (where the number 1 represents stability and 0 represents instability) that the optimal enforcement mechanism enhances coalition's stability as compared to the case without enforcement, but the extent of the stability enhancement is quite limited. Specifically, as shown in Tables 5.1 and 5.3 the coalition with three members can be stabilised as the largest stable coalition among the three model specifications if the fine function is linear or the monitoring probability is concave in monitoring costs. However, only the coalition with two members is stable when the monitoring function is linear and the fine function is quadratic. The result of the stability is reasonable by comparing signatories' abatement and payoffs under three model specifications. It can be seen from Tables 5.1 to 5.3 that signatories' abatement levels and the resulting payoffs are lowest under the model with concave monitoring probability function compared to the results in other two specified models, where achievements in abatement and payoffs under the model with quadratic fine function are the largest and signatories in the model with both linear fine and monitoring probability functions abate and gain less than that. As higher abatement level implies stronger free-riding incentives and vice versa, the stable coalition under the quadratic fine is the smallest but achieves the most.

The comparison of monitoring costs between the three model specifications shows that lower monitoring costs are required under the quadratic fine function compared to the one in other two models with linear fine form. This result confirms our analysis that under the enforcement mechanism with quadratic fine function, both of monitoring costs and target can be used as effective enforcement instruments. Hence, monitoring costs can be saved by setting higher target that results in heavier punishment. This counts for higher payoffs under the quadratic fine function as compared to the ones achieved under the linear fine function. By comparing the monitoring costs and the incentivised signatories' abatement level (see row 6 and 7 in Tables 5.1-5.3), it can be seen that in contrast to the linear monitoring probability function, the marginal abatement under the concave monitoring probability function is the lowest. This is due to the marginal monitoring probability being decreasing in monitoring expenditures.

Table 5.1. Payoffs and Stability for the basic model with a linear fine and a linear monitoring probability functions

	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$	$s=7$	$s=8$	$s=9$	$s=10$
$V_i^S(s)$	990.50	1130.50	1370.50	1710.50	2150.50	2690.50	3330.50	4070.50	4910.50
$V_i^{NS}(s-1)$	940.00	1130.00	1520.00	2110.00	2900.00	3890.00	5080.00	6470.00	8060.00
$V_i^{NS}(s)$	1130.00	1520.00	2110.00	2900.00	3890.00	5080.00	6470.00	8060.00	---
$V_i^S(s+1)$	1130.50	1370.50	1710.50	2150.50	2690.50	3330.50	4070.50	4910.50	---
m^*	18.00	57.00	116.00	195.00	294.00	413.00	552.00	711.00	890.00
q_i^*	19.00	29.00	39.00	49.00	59.00	69.00	79.00	89.00	99.00
Internal stability	1	1	0	0	0	0	0	0	0
External stability	0	1	1	1	1	1	1	1	---

Table 5.2a. Payoffs and Stability for the model with a quadratic fine function: for the case $q_i^* = \frac{sb}{a}$

	s = 2	s = 3	s = 4	s = 5	s = 6	s = 7	s = 8	s = 9	s = 10
$V_i^S(s)$	999.89	1149.75	1399.57	1749.33	2199.00	2748.50	3397.67	4146.00	4991.00
$V_i^{NS}(s-1)$	950.00	1150.00	1550.00	2150.00	2950.00	3950.00	5150.00	6550.00	8150.00
$V_i^{NS}(s)$	1150.00	1550.00	2150.00	2950.00	3950.00	5150.00	6550.00	8150.00	---
$V_i^S(s+1)$	1149.75	1399.57	1749.33	2199.00	2748.50	3397.67	4146.00	4991.00	---
m^*	0.22	0.75	1.71	3.33	6.00	10.50	18.67	36.00	90.00
q_i^*	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00
Internal stability	1	0	0	0	0	0	0	0	0
External stability	1	1	1	1	1	1	1	1	---

Table 5.2b. Payoffs and Stability for the model with a quadratic fine function: for the case $q_i^* < \frac{sb}{a}$

	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$	$s=7$	$s=8$	$s=9$	$s=10$
$V_i^S(s)$	999.89	1149.750	1399.57	1749.33	2199.00	2748.50	3397.67	4146.03	4991.42
$V_i^{NS}(s-1)$	949.90	1149.753	1549.53	2149.18	2948.61	3947.60	5145.64	6541.18	8128.04
$V_i^{NS}(s)$	1149.753	1549.53	2149.18	2948.61	3947.60	5145.64	6541.18	8128.04	---
$V_i^S(s+1)$	1149.750	1399.57	1749.33	2199.00	2748.50	3397.67	4146.03	4991.42	---
m^*	0.22	0.75	1.71	3.33	5.99	10.47	18.57	35.46	82.17
q_i^*	19.99	29.98	39.98	49.97	59.96	69.94	79.89	89.76	99.15
Internal stability	1	0	0	0	0	0	0	0	0
External stability	1	1	1	1	1	1	1	1	---

Table 5.3. Payoffs and Stability for the model with a linear fine and a concave monitoring probability functions

	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$	$s=7$	$s=8$	$s=9$	$s=10$
$V_i^S(s)$	989.63	1126.18	1359.11	1686.88	2106.96	2615.28	3205.26	3867.02	4587.72
$V_i^{NS}(s-1)$	940.19	1125.96	1504.84	2072.69	2822.23	3740.33	4803.54	5972.49	7191.77
$V_i^{NS}(s)$	1125.96	1504.84	2072.69	2822.23	3740.33	4803.54	5972.49	7191.77	---
$V_i^S(s+1)$	1126.18	1359.11	1686.88	2106.96	2615.28	3205.26	3867.02	4587.72	---
m^*	19.29	68.07	156.08	299.29	521.61	857.31	1394.44	2036.67	2930.03
q_i^*	18.80	28.49	38.07	47.44	56.51	65.05	72.78	79.35	84.55
Internal stability	1	1	0	0	0	0	0	0	0
External stability	0	1	1	1	1	1	1	1	---

5.4 Conclusion

Considering the free-riding incentives in the process of compliance with abatement commitments, individual signatory's strategic compliance level is conditional on the design of enforcement mechanism. As a result, signatories' payoffs and participation incentives in an ICA can also be changed. In this paper we use a combination of analytical and numerical analyses to explore the design of the optimal enforcement mechanism in self-enforcing ICAs and also to illustrate the effect of the optimally designed enforcement on signatories' compliance levels and the stability of ICAs.

A coalition decides on the optimal enforcement policy with respect to the choices of monitoring expenditure and the abatement target by maximising joint welfare. Under such an optimal enforcement mechanism, the coalitional first best described by Samuelson rule can be achieved when the fine imposed on defectors are marginally increasing in defection levels and meanwhile the marginal monitoring probability is constant in monitoring expenditure. Otherwise, a second-best in which individual signatory's abatement level is lower than the coalitional best level is achieved. Because monitoring is costly, full compliance under the optimal enforcement will usually not be enforced.

In our model set up with a quadratic fine function even the coalitional best abatement level is partial compliance. The target is set strategically to overly strict levels because setting a high target implies higher marginal fines and the target is a costless instrument to incentivise abatement. In this case full compliance should not be the objective of the enforcement policy design.

The stability of ICAs is improved by the optimal enforcement mechanism but to a very limited degree. Under a linear fine function, the stability of a coalition depends on enforcement parameters. Our analytical results show that defection costs become high if the parameters representing the severity of the punishment policy and the monitoring efficiency are taking high values. Large defection costs have an adverse effect on incentives to join a coalition and such that only the smaller coalition can be stabilised, the vice versa. This is in line with the result of McEvoy and Stranlund (2009), who argue that high values of enforcement parameters decrease the enforcement cost, such that only lower participation level is needed to increase coalition's abatement level. A narrow but deep coalition can be formed when the fine is increasing in the degree of the defection. Under the quadratic fine function the high abatement level and high resulting payoffs also cause high free-riding incentives. A slightly larger coalition can be sustained under the concave monitoring probability because the low abatement improvement, as compared to the enforcement with quadratic fine function.

In conclusion, our study shows that an optimal designed enforcement mechanism can increase the mitigation effectiveness of ICAs through inducing higher compliance levels, especially if the penalty for defectors is marginally increasing with the defection degree. Because free-riding incentives increase with the coalitional abatement level, broad and deep coalitions cannot be stabilised under a costly enforcement mechanism. However, this stability result could be different if the players are asymmetric in marginal abatement costs and benefits. This is one direction to extend our research. For asymmetric players, the optimal enforcement mechanism could result in the mixture of full and partial compliance in a coalition, thus signatories' incentives to join an ICA varies among signatories.

5.5 Appendix

Supposing the constraint on target setting is non-binding, then the shadow value of the constraint is $\gamma_i = 0$. Hence, Eq. (5.7b) can be rewritten as

$$\sum_{j \in S} \frac{\partial B_j}{\partial q^*} = \frac{\partial C_i}{\partial q_i^*}. \qquad (5.A1)$$

which is Samuelson's rule. By replacing $\sum_{j \in S} \frac{\partial B_j}{\partial q^*}$ in Eq. (5.7a) by $\frac{\partial C_i}{\partial q_i^*}$, the following equation is obtained:

$$\frac{\partial C_i}{\partial q_i^*} \sum_{j \in S} \frac{\partial q_j^*}{\partial m} = \sum_{j \in S} \left(\frac{\partial C_j}{\partial q_j^*} \frac{\partial q_j^*}{\partial m}\right) + 1. \qquad (5.A2)$$

Clearly, Eq. (5.A2) cannot hold in general implying $\gamma_i > 0$.

Chapter 6

Discussion and conclusions

Global cooperation among sovereign countries is an efficient way to mitigate greenhouse gas (GHG) emissions, however, it is difficult to realise due to the public good property of GHG emissions mitigation. This thesis investigates the formation and effectiveness of international climate agreements (ICAs) by exploring the effects of carbon trade, asymmetric bargaining powers and enforcement involving costly monitoring and punishment on improving the stability and effectiveness of ICAs. I use both game theory and numerical modelling to analyse these mechanisms. The analysis and results of this thesis provide insights into the design of current and future ICAs for improving the stability and effectiveness.

6.1 Answers to research questions

- What is the impact of a carbon market on regional incentives to join an ICA for GHG emissions mitigation when the carbon market is established independently of this agreement?

Carbon trade within a coalition for GHG emissions mitigation can improve incentives to join (Altamirano-Cabrera & Finus 2006). To investigate the impact of an independent carbon market that is open to all on participation incentives and performances of a climate coalition for GHG emissions mitigation, a four-stage coalition formation game is formulated in Chapter 2. In this four-stage game, the equilibrium coalition structure implies the simultaneous stability of the climate coalition and the carbon market. Being different from a single climate mitigation coalition, the presence of a carbon market enlarges the space of players' strategic choices and therefore changes incentives to join the mitigation coalition.

I firstly analyse the game theoretically by backward induction. Due to the complexity of a two-coalition structure, the analysis is based on different memberships of players. At stage 4, the singleton non-traders who are outside of both the carbon market and the mitigation coalition, choose their optimal abatement levels by equating their marginal benefits with their marginal costs of mitigation. For carbon traders who are members of the carbon market, it is optimal to take the abatement level that equates marginal mitigation costs with the permit price. For the mitigation coalition members who choose to be outside of the carbon market, they choose the optimal abatement level based on Samuelson's rule. Stage 3 only involves the behaviours of the carbon traders, i.e. their optimal allowance choices. I apply the endogenous allowance choice scheme developed by Helm (2003) to the strategic allowance choice at this stage, such that the solution for choosing the optimal allowance level can be derived. The stability check of the carbon market and the ICA at stage 1 and 2 cannot be realised through the theoretical analysis without specified functional forms. To gain insights into the stability and performance of this four-stage coalition formation game, I employ a modified version of STACO model where 12 world regions are aggregated into 7 regions and the related parameters are also recalibrated.

Results from the numerical analysis show that impacts of an independent carbon market on the mitigation coalition are complex. Firstly, with respect to the levels of stability and performance, this carbon market exerts a negative impact on the mitigation coalition. With the presence of the carbon market, the number of coalitions that can be formed decreases. At the same time, the best-performing mitigation coalition under the two-coalition structure generates lower global payoffs and abatement than the best-performing one in the absence of a carbon market. Secondly, for a given stable coalition, the presence of a carbon market can help to improve global abatement and payoffs. Global abatement increases with the presence of the carbon market, which involves more regions to abate. The improvement in global payoffs is driven by the positive externality of increased global abatement. Thirdly, with a carbon market, some instable coalitions under the one-coalition structure become stable. This conclusion especially applies to the internally stable but externally instable coalitions, which become fully stable coalitions with the presence of the carbon market. The intuition behind this is that the independent carbon market provides an additional opportunity for some non-signatory countries to increase mitigation efforts through carbon trade. This induces a Pareto improvement on abatement levels and profitability compared to the one-coalition structure.

Although this independent carbon market is open to both the mitigation coalition and its outsiders, the numerical analysis shows that in equilibrium members of the mitigation coalition are not willing to join the carbon market. As shown by the theoretical analysis, the mitigation coalition is more likely to be a permit buyer if the coalition joins the carbon market. Hence, the mitigation coalition receives negative transfers from buying emission permits. For the mitigation coalition, the free-riding benefits surpass the gain from joining carbon trade.

In conclusion, even with this dual-agreement system, the structure of coalition formation is still a partial participation with free-riding outsiders. More alternative strategic options offered by carbon trade destabilise the most efficient mitigation coalitions that are stable in the system with a single mitigation agreement. However, it could be beneficial to some mitigation coalitions by expanding the single coalition structure to improve the efficiency of some stable, but suboptimal ICAs.

- How does an individual allowance choice constraint impact the stability and effectiveness of a carbon trade agreement?

Without constraints on allowance choices, a carbon market would not be very effective (Helm 2003). Moreover, participation incentives that stem from unlimited allowance choices could also dampen the stability of a carbon market with open membership. Therefore, imposing a constraint on allowance choices might not only mitigate the hot-air effect but can also help to stabilise a carbon market by avoiding excessive participation of potential sellers. In Chapter 3 I study the impact of a constraint on individual allowance choices on the stability and performance of a carbon market with open membership. The carbon market is formulated as a coalition focusing on carbon trade, and the analysis is based on a standard two-stage coalition formation model. At stage one all players make decisions on signing or not signing a carbon trade agreement, which obliges all signatories to accept a constraint on the allowance choice. The constraint imposed on individual carbon traders is modelled as a fraction of each trader's business-as-usual (BAU) emissions. In the second stage, carbon traders simultaneously choose emission allowances subject to exogenous constraints, and trade and decide their after-trade abatement assuming the abatement of all non-traders is given. Thus, the outsiders choose their abatement by taking others' abatement levels as given.

The sub-game perfect Nash equilibria of this two-stage game are derived by backward induction. The results from theoretical analysis show that the constraint imposed on allowance choices is not binding when the constraint is lax. Under non-binding constraints, allowance choices and the carbon price in the constrained carbon market where the shadow value of carbon is zero, are identical to the carbon market without constraints. By increasing the strictness of the constraint, the constraint becomes binding and the carbon price / shadow price of carbon increases. Under binding constraints, carbon traders choose the optimal allowances that are equal to this imposed 'cap'. Because of the positive shadow value of carbon allowances when constraints are binding, the gains of all carbon traders decrease with the tightening of the constraint on allowance choices. As a result, there is no participation incentive and thus no carbon market will emerge when the constraint is too strict. This indicates that the condition for the internal stability is violated.

To illustrate the consequences of imposing constraints on individual allowance choices more explicitly, a numerical analysis based on this two-stage game is also implemented by employing the STACO model. As a base scenario, a numerical analysis for an unconstrained carbon market is performed. The numerical results for the base scenario show that the hot-air effect takes place and the stability of a carbon market with open membership is dampened if the allowances can be chosen arbitrarily. By imposing a constraint on individual allowance choices, the numerical analysis shows that the hot-air effect can be alleviated and even eliminated by increasing the strictness of the constraint on allowance choices. At the same time, this constraint also has a positive effect on the stability of the carbon market in the sense that multiple equilibrium carbon markets can be formed under different constraints. This result confirms the theoretical analysis with respect to that no stable carbon market can be found when the constraint is set too strict. As to the impact on mitigation effectiveness of the carbon market, global abatement and welfare are enhanced due to the limited allowance choices. However, the improvement of the global abatement and welfare reaches the highest levels when the constraint is set at a moderate level.

With a sensitivity analysis, the impact of the baseline setting on the stability and effectiveness of constrained carbon markets is also analysed numerically. Instead of the BAU baseline for the constraint on individual allowance choices, the non-cooperative Nash emissions levels become the new baseline for the constraint. Numerical results indicate that larger and more effective carbon markets can be formed under the Nash baseline for the constraint as compared to the BAU baseline. The improved stability and effectiveness of stable carbon markets with Nash baseline are related to the binding allowance choices of all carbon traders, which are found in all carbon markets from our numerical results. In particular, a carbon market with full participation can be sustained when the upper bound of individual allowance choices is the Nash-emissions level. The intuition is that each region can only improve but never lose upon its Nash payoff by joining the carbon market.

In conclusion, under a carbon market with open membership, the stability and the membership of an international carbon market can be increased by imposing a constraint on allowance choices. When tightening the constraint '*broad but shallow*' agreements are replaced by '*narrow but deep*' ones. Imposing constraints on allowance choices can improve the stability and enlarge the scale of an international carbon market to a limited degree, but cannot overcome the free-riding incentives.

- What is the impact of using the Nash bargaining solution for distributing coalitional gains under different sets of bargaining weights on the stability and effectiveness of international climate agreements?

Chapter 4 examines how the formation of ICAs and their mitigation efficiency are impacted by the use of the Nash bargaining solution (NBS), which is affected by negotiators' unequal bargaining power. In a Nash bargaining game with heterogeneous players, bargaining powers are unequal and may be driven by different characteristics of negotiators. Potential reasons that could result in differences of negotiators' bargaining power in international climate negotiations are firstly discussed and reviewed in Chapter 4. Based on the discussion of these potential factors, five different determinants that may change negotiators' bargaining power are identified and then are used for the quantification of negotiators' bargaining weights. These five determinants are: (i) discount factor; (ii) abatement efforts; (iii) abatement costs; (iv) climate change damages; (v) economic power (in terms of GDP). Furthermore, the outside options of players, which also can impact the bargaining solution, are introduced. It is argued that the Nash bargaining solution with outside options (see Muthoo 1999) falls in the class of optimal sharing rules described by Carraro et al. (2006) and Weikard (2009).

To compare and examine the impact of the NBS with different sets of unequal bargaining weights on incentives to cooperate, international climate negotiations are modelled as a two-stage cartel formation game. In this model, gains from climate mitigation cooperation are distributed among signatories by applying the NBS with different sets of bargaining weights. Based on the two-stage game, a numerical analysis is implemented in Chapter 4. The numerical results show that multiple equilibrium climate coalitions can emerge from the NBS. The number and size of stable coalitions vary with different sets of bargaining weights. Moreover, it turns out that the success of international climate agreements depends on the set of bargaining weights in climate negotiations. Some sets of bargaining weights generate more successful coalitions in terms of welfare and abatement than others. For example, among five sets of asymmetric bargaining weights, the one determined by negotiators' economic power can facilitate a climate coalition that comprises two of the largest emitters (CHN and USA) jointly with two other regions.

Numerical results in Chapter 4 also show that the NBS with outside options is more conducive to ICAs as compared to the bargaining solution without considering outside options. The analysis on the NBS with outside options provides a rationale for the use of optimal sharing rules: they result from a NBS with outside options.

It can be concluded from the study of Chapter 4 that negotiators' bargaining power in international climate negotiations can be interpreted as and quantified with different sets of bargaining weights. Accordingly, Nash bargaining solutions in terms of the results of coalitional gains allocation among members are differentiated by applying different sets of bargaining weights. By using the NBS as a transfer scheme, the participation incentives and performances of ICAs can be improved as compared to agreements that do not redistribute gains from cooperation, but its capacity to overcome free-riding incentives is limited. However, if Nash bargaining accounts for outside options of players, larger stable coalitions and higher global abatement levels can be achieved.

- How can an optimal enforcement mechanism for an ICA be designed, and what is the impact of an optimally designed enforcement mechanism on participation and compliance?

The success of an international climate agreement is not only relying on the participation but also on the implementation of abatement commitments by signatories. To improve the implementation of abatement commitments, an enforcement mechanism is required to incentivise signatories to comply with their commitments. Chapter 5 investigates the optimal enforcement policy with respect to the optimal abatement target and monitoring investment for an ICA. Furthermore, I analyse how an optimally designed enforcement mechanism changes incentives to participate and comply in a coalition formation game. The model in Chapter 5 is based on the enforcement mechanism introduced by McEvoy and Stranlund (2009, 2010) but is extended by considering partial compliance. In this model with partial compliance, the optimal compliance level of signatories is determined endogenously and strategically based on the monitoring expenditure and abatement targets. The model in this paper is formulated as a four-stage coalition formation game. At the first stage, countries make their membership choices. At the second stage, coalition members jointly and simultaneously fix the abatement target and monitoring expenditures that determine the probability of monitoring. At the third stage, countries choose their abatement levels independently. Singletons just choose a best response while signatories must take the expected punishment into account, in case they would not comply. At the final stage, signatories' abatement levels are randomly monitored and fines have to be paid if non-compliance is detected.

The four-stage game is solved analytically by backward induction. The analysis gives the following findings on the optimal enforcement design and signatories' compliance levels. Firstly, a signatory chooses the optimal abatement level by equalising marginal net gains of abatement with the marginal expected fine. This implies that Samuelson's rule which determines the coalitional best level of abatement can be satisfied if the marginal expected fine makes up for the externality of each signatory's abatement. Secondly, signatories to an ICA decide cooperatively on the enforcement mechanism with respect to the abatement target and monitoring expenditures. In equilibrium, monitoring expenditures are decided so that coalitional marginal net gains of the increased abatement due to one unit increase in monitoring costs is equal to the marginal monitoring costs. With a constraint to the target setting, the target is optimally chosen at the upper bound value of the constraint that is the BAU emissions level. Note that, this equilibrium condition for the optimal target setting only applies to the case where the target can play a role in inducing higher abatement levels, for example when the marginal fine is increasing in defection levels. Under such an optimal target, signatories are incentivised to choose an abatement level that is either equal to or smaller than the coalitional best one obtained from Samuelson's rule. This result has two implications: firstly, under the optimal enforcement mechanism, the target and the monitoring are used as two instruments to increase signatories' compliance level. However, setting a higher target is costless while increasing the monitoring probability is not. Therefore, it is always better to choose the maximal target level when implementing Samuelson's rule. But note that the target is not an effective enforcement instrument when the marginal fine is a constant value. Secondly, setting higher targets can increase the abatement, but it is not optimal for a coalition to abate more than the Samuelson level. This is true since under the Samuelson condition, each coalition member makes the abatement choice by internalising the external effects on all other signatories.

To explore the impact of the design feature (i.e. the penalty structure and monitoring technology) of enforcement mechanism in ICAs, I apply the analytical solutions to three cases with specific functions. Several results can be obtained from the analysis of the three specified models. Firstly, the optimal enforcement policy induces partial compliance, but the level of which varies under different functional forms of fine and monitoring probability. Under the enforcement with quadratic fine and linear monitoring probability functions, signatories' compliance level can be induced to the most since both the target and monitoring costs are effective in incentivising signatories to take the coalitional best abatement level. Under the enforcement mechanism with linear fine and linear monitoring probability functions, signatories' compliance level can be incentivised only by monitoring efforts but not by the target due to the constant marginal fine. Hence, due to costly monitoring, signatories' optimal abatement level induced by the optimal monitoring costs that are chosen by the coalition is lower than the coalitional best level. Monitoring technology improvement can have a positive effect on incentivising signatories' compliance level. As shown in our numerical results, signatories' compliance level is the lowest when the marginal monitoring probability is decreasing in monitoring costs, as compared to the results from the constant marginal monitoring probability.

Secondly, due to different compliance levels, the payoffs and participation incentives in an ICA are also different under different functional forms of the fine and monitoring probability. In the case of a linear fine function, the enforcement parameters representing the productivity of monitoring expenditure and the severity of the punishment respectively affect the stability of an ICA. The intuition is that higher values of enforcement parameters imply higher defection costs and thereby reduces the incentives to join an ICA. This result also implies that the grand coalition can be stabilised if enforcement parameters have low values under the enforcement mechanism with linear fine functions. Due to the implicit analytical solutions, the stability results under the quadratic fine and the concave monitoring probability functions cannot be drawn explicitly. By using a numerical example, our results illustrate that the stable coalition under the enforcement mechanism with the quadratic fine and linear monitoring probability functions contains the fewest members, while a slightly larger coalition can be stabilised if the fine function is linear or the monitoring probability function is concave. The reason is that the abatement and payoffs under a quadratic fine function can achieve the highest levels, which also generates the strongest free-riding incentives. However, the enforcement mechanism with concave monitoring probability has the lowest effectiveness in incentivising signatories' abatement levels. Hence, free-riding incentives are low under the enforcement with linear fine or concave monitoring probability. This stability results implies that under the optimal enforcement policy, an ICA can either be formed as a narrow and deep one or as a broad but shallow one.

6.2 General discussion and conclusions

6.2.1 Modelling conclusions

Game-theoretical modelling is a useful and essential instrument to study the problem of international cooperation on climate change mitigation. Game theory can facilitate the analysis of countries' strategic interactions and derivation of optimal strategies. In the context of climate change cooperation, non-cooperative game theory typically models the formation of ICAs as a two-stage game: the first stage is the membership choice, and the second stage involves the optimal abatement choice. To study the impact of a carbon market that is open to all on the stability and effectiveness of a climate mitigation coalition in Chapter 2, I extend the typical two-stage game to a more complex four-stage game. In this four-stage game, there are two coalitions: a carbon market and a climate coalition for GHG emissions mitigation. These coalitions are formed sequentially at stages 1 and 2. At stage 3, market participants make their decisions on allowances choice and in the final stage all players choose abatement levels. Through this four-stage coalition formation model, I find that the carbon market has an adverse effect on countries' incentives to participate in the mitigation coalition. This is explained by the alternative strategic options offered by the carbon market, which enlarge each player's strategy choice space. As a result, players who intend to join the mitigation coalition for collective gains can also be attracted by carbon trade for low marginal mitigation costs, or opt to be a free-rider because of the increased abatement level resulting from carbon trade. As a consequence, the size of the mitigation coalition is reduced compared to the case without the presence of the carbon market. Furthermore, modelling the carbon market as a one that is open to all enables us to analyse the mitigation coalition's incentive to join carbon trade. The results in Chapter 2 show that the carbon market is not attractive for the mitigation coalition, which prefers to be a market outsider due to the free-riding incentives from the increased mitigation level.

The carbon market studied in Chapter 2 is unconstrained, which results in hot-air and excessive participation incentives that may result in instability of the carbon market. Hence, I impose an exogenous constraint on individual carbon trader' allowance choices and study the effect of this constraint on the stability of the carbon market in Chapter 3. I employ a standard two-stage coalition formation game, where all players simultaneously make decisions on their market participations at stage 1, and signatories make their allowance choices and then trade at stage 2. The results in Chapter 3 shows that both of the stability and the membership of an international carbon market can be increased by imposing a constraint on allowance choices. Moreover, when tightening the constraint '*broad but shallow*' agreements are replaced by '*narrow but deep*' ones.

To compare and examine the impact of the Nash bargaining solutions (NBS) with different sets of unequal bargaining weights on incentives to cooperate on climate change mitigation, I formulate a standard two-stage cartel formation game in Chapter 4. All players make their membership choices at stage one and abatement targets are set cooperatively by signatories at stage two. In this game, signatories' individual payoffs are determined by applying the NBS with a given set of bargaining weights, which are based on signatories' asymmetric bargaining power. The conclusion in Chapter 4 is that applying the NBS without outside options to distribute coalitional gains, players' incentives to participate and abate can be increased but to a limited degree. However, by considering outside options, the stability of an ICA can be improved to a further degree. The reason is that the bargaining outcome falls into the class of optimal sharing rules when outside options are considered. Hence, the study in Chapter 4 provides a rationale for the use of optimal sharing rules: the optimal sharing rules result from a NBS with outside options.

By formulating a four-stage coalition formation model in Chapter 5, I investigate impacts of an optimally designed enforcement mechanism on signatories' compliance levels and the stability of an ICA. I conclude that the optimal enforcement induces partial compliance. However, compliance levels are different under different functional forms of fine and monitoring probability. Under an enforcement mechanism with fines for defectors are increasing in defection levels and the marginal monitoring probability is constant, both of monitoring costs and the abatement target are effective in incentivising higher compliance levels. As a result, the coalition's first-best abatement level can be induced. However, when both of the marginal fine and monitoring probability are constant, only monitoring costs are the effective enforcement instrument and thus signatories' compliance levels are lower than the one under quadratic fine function. Monitoring probability matters for the compliance level. A concave monitoring probability function is less effective in inducing higher abatement. A lower compliance level results when the marginal monitoring probability is decreasing in monitoring costs. In terms of the stability effect, the stability of ICAs can be improved by the optimal enforcement mechanism but with very limited degrees.

6.2.2 Policy conclusions

Being adopted on 12 December 2015 and entering into force on 4 November 2016, the Paris agreement has been a highly expected institutional instrument governing global GHG emissions mitigation. However, there are still challenges facing the Paris agreement, for example, instability arising from member countries' dropping out due to economic and political reasons, low mitigation effectiveness or poor implementation of ambitious targets. By focusing on these common challenges that can be faced by all ICAs, this thesis studies the institutional dimension of the global climate change problem and aims to increase the degree of success of ICAs by employing economic and institutional instruments. Results of this thesis can be interpreted and lead to the following policy implications.

First, establishing a carbon market independently of an ICA is ineffective to increase cooperation but can increase global abatement level by involving more countries to abate. In reality an example of such a carbon market can be referred to the CDM, under which uncommitted developing countries under the Kyoto Protocol can indirectly reduce GHG emissions by allowing the committed countries to implement emission-reduction projects in their lands. The results in Chapter 2 show that a carbon market that is established outside of an ICA cannot facilitate the enlargement of an ICA. The intuition is that each country's strategy choices are increased due to the presence of a carbon market as compared to a single ICA, therefore countries seek cheaper abatement options not only by joining an ICA but also by joining a carbon market. However, the global abatement level is increased by the presence of the carbon market because outsiders of the ICA can become signatories and do more abatement than the non-cooperative Nash level by joining carbon trading. For policy makers, this could be a solution to increase the mitigation effectiveness of partial climate coalitions given that a global climate agreement is difficult to reach.

Second, imposing a constraint with the baseline of non-cooperative Nash emissions level on individual countries' allowance choices can help to increase participation in and mitigation effectiveness of an international carbon market. This policy implication can be interpreted in two ways. One refers to the importance of designing constraints on carbon trader countries' emission allowance choices in an international carbon market. This constraint can increase the mitigation effectiveness of a carbon market by limiting the amount of the emission allowances, but can also increase the stability of a carbon market by regulating excessive participation incentives. As shown by the analysis of Chapter 3, without a constraint on allowance choices it is difficult for a carbon market with open membership to satisfy external stability, because it is easily destabilized by market entrants who seek for revenues from carbon sale. The lower a region's marginal damages of GHG emissions, the more allowances will be chosen. Excessive allowance choices result in an inefficient carbon market. By imposing constraints on allowance choices, the hot-air effect can be eliminated. Moreover, the stability of a carbon market can be enhanced in a way that the incentives of obtaining revenues from carbon sale can be controlled through limiting the allowance choices. Hence, the policy with respect to imposing a constraint on individual allowance choices can be considered for a coalition of international carbon trade.

The second interpretation concerns the choice of a baseline for the constraint on allowance choices. The IPCC Fourth Assessment Report (AR4) (IPCC 2007) defines baseline as "the reference for measurable quantities from which an alternative outcome can be measured, e.g. a non-intervention scenario is used as a reference in the analysis of intervention scenarios." Accordingly, the baseline levels used in the current climate change negotiations, for example the 1990 year's GHG emissions level or the pre-industrial level, can be taken as the BAU emissions level that is projected without policy intervention. However, by reducing damages from and the vulnerability to climate change, currently most countries in the world make domestic policies by taking the impact of climate change into account. Especially after the Copenhagen Conference, developing countries also announced their mitigation actions. This kind of policies, which are made based on individual welfare maximisation and aim to internalise the negative effects of climate change on their own, results in the so-called non-cooperative Nash equilibrium emissions. Hence, the baseline in current international climate negotiations should be revised by taking the Nash emissions level. One of the findings in Chapter 3 suggests that altering the constraint baseline from the BAU level to the non-cooperative Nash emissions level could ease negotiations. Under the Nash baseline, not only the global welfare and mitigation effectiveness of a carbon market are improved but also the grand carbon market might be formed. This is because as compared to the non-cooperative scenario, each region can only improve upon its Nash payoffs by joining the carbon market.

Third, in international climate negotiations countries' bargaining power could be determined by various factors. This provides a reference for policy makers with respect to the elaboration on coalitional surplus sharing rules that can improve the mitigation effectiveness of ICAs. For example, analysis of the bargaining power of key GHG emitters could be used to induce the participation of key emitters. In the model of Chapter 4, the distribution of gains is the outcome of a bargaining process with unequal bargaining power and, thus, sharing weights of collective gains are determined by bargaining power. The numerical results in Chapter 4 suggest that different sets of bargaining weights quantified from determinants of bargaining power affect the formation and effectiveness of ICAs to different degrees. For example, the bargaining weights determined by negotiators' economic power can facilitate the best-performing climate coalition, which includes the world largest GHG emitters, the USA and China. The numerical results from Chapter 4 also show that in the negotiation process multiple determinants of bargaining power play a role. The intuition is that one country's incentives to cooperate on GHG mitigation are impacted in a complex way by factors that are related to abatement options, climate change vulnerability and economic power. Hence, the sharing weights can also be derived from the combination of multiple determinants of bargaining power.

Lastly, an enforcement mechanism with fines for non-compliance should be designed and adopted to increase compliance of signatories of ICAs. In the Paris agreement, countries' abatement targets, which is set by themselves and called the "nationally determined contributions" (NDCs), are not binding. Hence, it is necessary to reduce the risk of noncompliance by establishing an enforcement mechanism consisting of monitoring and punishment of defectors. As shown by results in Chapter 5, signatories' compliance levels can be improved by an optimal enforcement mechanism. Especially when fines imposed are marginally increasing in the defection level, the coalitional best abatement level can be induced. This is because signatories can be incentivised to increase their compliance levels by both higher monitoring costs and higher targets under the enforcement mechanism with quadratic fine function. Due to the lack of supranational agency in international climate agreements, member countries' compliance with mitigation targets is hardly to be monitored and then be enforced. However, assuring compliance with abatement commitments is crucial for the effectiveness of ICAs. Therefore, designing an optimal enforcement mechanism and successfully implementing it are urgent issues facing the Paris or post-Paris international climate agreements. The study in Chapter 5 can be considered as a reference for that.

6.3 Limitations and suggestions for further research

The numerical analysis is implemented in Chapter 2 to 4 by employing a numerical (STACO) model. Therefore, results and conclusions from these three chapters are restrictive with regard to the specific model features. As to the game setting, STACO is based on the concept of cartel formation game. Hence, the game theoretical analyses in Chapter 2 to 4 all focus on the structure of one international climate coalition, which excludes the case of multiple coalitions. Moreover, the stage games in Chapter 2 to 4 are all static and do not take the uncertainty into account. All players make their decisions on membership and abatement level in the pre-period and then decisions last for all future periods. This static game setting ignores the impact of the current decisions on the ones in the future, hence it is reasonably for decision makers to change the strategies for coalition participation and abatement magnitude over time. Climate change projections are highly uncertain, which requires to use the probabilistic methods to estimate and tackle the climate change problem (Weitzman 2009; Millner et al. 2013). Therefore, for future research, it would be interesting to study the formation of ICAs and the optimal mitigation problem by employing a stochastic-dynamic control model. As an example, Kolstad (2007) introduces uncertainty and learning in environmental costs and benefits to a standard formation model of international environmental agreements (IEAS). He finds that systematic uncertainty designed in his model has a negative effect on the size of an IEA. The effect of learning is ambiguous.

In Chapter 3, the constraint on allowance choices is modelled as an exogenous parameter that is common for all heterogeneous carbon traders. Due to different marginal abatement costs and benefits, it would be more realistic to set conditional constraints for different carbon traders. The exogenous setting of the constraint makes the analysis more tractable, however the optimal constraint level under which a stable climate coalition can achieve the highest mitigation levels as compared to the ones under other constraint levels, cannot be determined theoretically for general conditions. Thus it would be valuable to extend the model to include a pre-negotiation stage where all players first determine their own constraint - possibly conditional on regional characteristics - and only after that the membership and allowance choices would follow. This analyse could provide an endogenous solution for the individual constraint level.

Chapter 4 argues that bargaining power of heterogeneous countries depends on some external characteristics of the country itself. However, if negotiating countries anticipate that their decisions (e.g. mitigation decisions) have an influence on future bargaining power, interesting strategic interactions may arise. One direction to extend our analysis is to study negotiators' strategic behaviour when bargaining power becomes an endogenous variable. This requires an extended dynamic game setting where pre-negotiations determine the negotiation protocol (Wangler et al. 2013) and thereby bargaining powers are relevant at the later stages of the game.

Chapter 5 studies the design of the optimal enforcement mechanism and its impact on the compliance level and stability of ICAs. Results show that because free-riding incentives increase with the coalitional abatement level, broad and deep coalitions cannot be stabilised under the optimal enforcement mechanism. However, this stability result is derived based on a model with symmetric players. The impact on stability of ICAs could be different if the players are characterised by asymmetric marginal abatement costs and benefits. This is one direction to extend our research. For asymmetric players, the optimal enforcement mechanism could result in the mixture of full and partial compliance in a coalition, thus signatories' incentives to join are different.

References

Altamirano-Cabrera, J.-C., Finus, M. (2006). Permit trading and stability of international climate agreements. Journal of Applied Economics, IX(1), 19-47.

Arguedas, C. (2005). Bargaining in environmental regulation revisited. Journal of Environmental Economics and Management, 50(2), 422–433.

Arguedas, C. (2008). To comply or not to comply? Pollution standard setting under costly monitoring and sanctioning. Environmental Resource and Economics, 41(2), 155-168.

Asheim, G. B., Froyn, C. B., Hovi, J., Menz, F. C. (2006). Regional versus global cooperation for climate control. Journal of Environmental Economics and Management, 51(1), 93-109.

Barrett, S. (1990). The problem of global environmental protection. Oxford Review of Economic Policy, 6(1), 68-79.

Barrett, S. (1992). Acceptable allocations of tradable carbon emission entitlements in a global warming treaty in: United Nations. Combating Global Warming–Study on a global system of tradable carbon emission entitlements, New York, 85-114.

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

Barrett, S. (1997). Heterogeneous international environmental agreements. In Carraro, C. (Ed.), International Environmental Agreements: Strategic Policy Issues. E. Elgar: Cheltenham.

Barrett, S. (1999). A theory of full international cooperation, Journal of Theoretical Politics, 11(4), 519–541.

Barrett, S. (2001). International cooperation for sale. European Economic Review, 45(10), 1835-1850.

Barrett, S. (2008). Climate treaties and the imperative of enforcement, Oxford Review of Economic Policy, 24 (2), 239–258.

Barrett, S., Stavins, R. (2003). Increasing participation and compliance in international climate change agreements. International Environmental Agreements: Politics, Law and Economics, 3(4), 349-376.

Becker, G. S.(1968). Crime and Punishment: An Economic Approach. Journal of Political Economy, 76, 169-217.

Behr, T., Witte, J.M., Hoxtell, W., Manzer, J. (2009). Towards a global carbon market? Potential and limits of carbon market integration, GPPI Policy Paper No.5, Berlin: Global Public Policy Institute.

Bernheim, B. D., Peleg, B., Whinston, M. D. (1987). Coalition-proof Nash equilibria I. concepts. Journal of Economic Theory, 42(1), 1–12.

Binmore, K., Osborne, M. J., Rubinstein, A. (1992). Noncooperative models of bargaining. In R.J. Aumann, S. Hart (Eds), Handbook of game theory with economic applications, 1, 179-225. North Holland: Amsterdam.

Binmore, K., Rubinstein, A., Wolinsky, A. (1986). The Nash bargaining solution in economic modelling. The RAND Journal of Economics, 17, 176-188.

Bohm, P., Larsen, B. (1994). Fairness in a tradeable-permit treaty for carbon emissions reductions in Europe and the former Soviet Union. Environmental and resource Economics, 4(3), 219-239.

Bosetti, V., Carraro, C., Galeotti, M., Massetti, E., Tavoni, M. (2006). WITCH-A world induced technical change hybrid model. University Ca'Foscari of Venice Economic Research Paper, No. 46/06. http://ssrn.com/abstract=948382 or http://dx.doi.org/10.2139/ssrn.948382 .

References

Botteon, M., Carraro, C. (1997). Burden-sharing and coalition stability in environmental negotiations with asymmetric countries. In C. Carraro (Ed.), International environmental negotiations: Strategic policy issues, 26-55. Edward Elgar: Cheltenham.

Brandt U.S. (2004). Unilateral actions, the case of international environmental problems. Resource and Energy Economics 26, 373–391.

Bréchet, T., Gerard, F., Tulkens, H. (2011). Efficiency vs. stability in climate coalitions: a conceptual and computational appraisal. The Energy Journal, 32(1), 49-76.

Buchanan, J.M. (1975). The limits of liberty. Between Anarchy and Leviathan. Chicago: University of Chicago Press.

Carbone, J.C., Helm, C., Rutherford, T.F. (2009). The case for international emission trade in the absence of cooperative climate policy. Journal of Environmental Economics and Management, 58, 266-280.

Carraro, C., Eyckmans, J., Finus, M. (2006). Optimal transfers and participation decisions in international environmental agreements. The Review of International Organizations, 1(4), 379-396.

Carraro, C., Siniscalco, D. (1993). Strategies for the international protection of the environment. Journal of Public Economics, 52(3), 309-328.

Carraro, C., Siniscalco, D. (1995). R&D cooperation and the stability of international environmental agreements (No. 1154). CEPR Discussion Papers.

Chander, P., Tulkens, H. (1995). A core-theoretic solution for the design of cooperative agreements on transfrontier pollution. International Tax and Public Finance, 2, 279-293.

Chander, P., Tulkens, H. (1997). The core of an economy with multilateral environmental externalities, International Journal of Game theory, 26 (3), 379-401.

Chwe, M. S. Y. (1994). Farsighted coalitional stability. Journal of Economic Theory, 63(2), 299-325.

Costantini, V., Sforna, G., Zoli, M. (2015). Interpreting bargaining strategies of developing countries in climate negotiations-A quantitative approach. SEEDS Working Paper.

Courtois, P., Tazdaït, T. (2014). Bargaining over a climate deal: deadline and delay. Annals of Operations Research, 220, 205–221.

D'Aspremont, C., Jaquemin, A., Gabszewicz, J.J., Weymark, J.A. (1983). On the stability of collusive price leadership. Canadian Journal of Economics, 16 (1), 17–25.

D'Aspremont, C.A., Gabsewicz, J.J. (1986). On the stability of collusion. In: Matthewson, G.F., Stiglitz, J.E. (Eds.), New Developments in the Analysis of Market Structure. MacMillan Press, New York.

Dellink, R. (2011). Drivers of stability of climate coalitions in the STACO model. Climate Change Economics, 2(2), 105-128.

Dellink, R.B., de Bruin, K., Nagashima, M., van Ierland, E.C., Urbina-Alonso, Y., Weikard, H.-P., Yu, S. (2015). STACO Technical Document 3: Model description and calibration of STACO 3. Wageningen School of Social Sciences, Working Paper No. 12, Wageningen University.

De Zeeuw, A. (2008). Dynamic effects on the stability of international environmental agreements. Journal of Environmental Economics and Management, 55(2), 163-174.

Edmonds, J., Wise, M., Barns, D. W. (1995). Carbon coalitions: the cost and effectiveness of energy agreements to alter trajectories of atmospheric carbon dioxide emissions. Energy Policy, 23(4), 309-335.

Eyckmans, J., Tulkens, H. (2003). Simulating coalitionally stable burden sharing agreements for the climate change problem. Resource and Energy Economics, 25(4), 299-327.

Eyckmans, J., Finus, M. (2006). Coalition formation in a global warming game: how the design of protocols affects the success of environmental treaty-making. Natural Resource Modelling, 19(3), 323-358.

Endres, A., Finus, M. (2002). Quotas may beat taxes in a global emission game. International Tax and Public Finance, 9, 687–707.

Fankhauser, S. (1995). Valuing climate change: The economics of the greenhouse. Earthscan, London.

Finus, M. (2003). Stability and design of international environmental agreements: the case of transboundary pollution. In: Folmer, H. and Tietenberg, T. (Eds.) The International Yearbook of Environmental and Resource Economics 2003/2004, 82-152. Edward Elgar: Cheltenham.

Finus, M. (2008). Game theoretical research on the design of international environmental agreements: insights, critical remarks, and future challenges. International Review of Environmental and Resource Economics, 2, 29-67.

Finus, M. (2008). The enforcement mechanisms of the Kyoto protocol: flawed or promising concepts? Letters in Spatial and Resource Sciences, 1, 13–25.

Finus, M., Altamirano-Cabrera, J.-C., and van Ierland, E.C. (2005). The effect of membership rules and voting schemes on the success of international climate agreements. Public Choice, 125, 95-127.

Finus, M., Rundshagen, B., (1998). Toward a positive theory of coalition formation and endogenous instrumental choice in global pollution control. Public Choice, 96, 145–186.

Finus, M., Saiz, M. E., Hendrix, E. M. (2009). An empirical test of new developments in coalition theory for the design of international environmental agreements. Environment and Development Economics, 14(1), 117-137.

Flachsland, C., Marschinski, R., Edenhofer, O. (2009). To link or not to link: benefits and disadvantages of linking cap-and-trade systems. Climate Policy, 9(4), 358-372.

Folmer, H., van Mouche, P. (2000). Transboundary pollution and international cooperation. International Yearbook of Environmental and Resource Economics 2000/2001, 231-266.

Grout, P. A. (1984). Investment and wages in the absence of binding contracts: A Nash bargaining approach. Econometrica: Journal of the Econometric Society, 449-460.

Han, Z., Ji, Z. J., Liu, K. J. (2005). Fair multiuser channel allocation for OFDMA networks using Nash bargaining solutions and coalitions. Communications, IEEE Transactions on Communications, 53(8), 1366-1376.

Helm, C. (2003). International emissions trading with endogenous allowance choices. Journal of Public Economics, 87(12), 2737–2747.

Hoel, M., 1992. International environmental conventions: the case of uniform reductions of emissions. Environmental and Resource Economics, 2(2), 141–159.

Hovi, J., Areklett, I. (2004). Enforcing the climate regime: game theory and the Marrakesh Accords. International Environmental Agreements: Politics, Law and Economics, 4, 1–26.

Hovi, J., Greaker, M., Hagem, C., Holtsmark, B. (2012). A credible compliance enforcement system for the climate regime. Climate Policy, 12, 741–754.

Intergovernmental Panel on Climate change (IPCC) (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Core Writing Team, R.K. Pachauri and L.A. Meyer (Eds.). IPCC, Geneva, Switzerland.

References

Intergovernmental Panel on Climate change (IPCC) (2007). Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Core Writing Team, Pachauri, R.K. and Reisinger, A. (Eds.). IPCC, Geneva, Switzerland.

Jackson, M., Moulin, H. (1992). Implementing a public project and distributing its cost. Journal of Economic Theory, 57(1), 125-140.

Larsen, B., Shah, A. (1994). Global tradeable carbon permits, participation incentives, and transfers. Oxford Economic Papers, 46(Supplement 1), 841-856.

Lange, A., Vogt, C. (2003). Cooperation in international environmental negotiations due to a preference for equity. Journal of Public Economics, 87(9), 2049-2067.

Kolstad, C. D. (2007). Systematic uncertainty in self-enforcing international environmental agreements. Journal of Environmental Economics and Management, 53(1), 68-79.

Leiby, P., Rubin, J. (2001). Intertemporal permit trading for the control of greenhouse gas emissions. Environmental and Resource Economics, 19(3), 229-256.

Lessmann, K., Edenhofer, O. (2011). Research cooperation and international standards in a model of coalition stability. Resource and Energy Economics, 33(1), 36-54.

Lessmann, K., Marschinski, R., Edenhofer, O. (2009). The effects of tariffs on coalition formation in a dynamic global warming game. Economic Modelling, 26(3), 641-649.

Lessmann, K., Marschinski, R., Finus, M., Kornek, U., Edenhofer, O. (2014). Emissions trading with non-signatories in a climate agreement—an analysis of coalition stability. The Manchester School, 82(S2), 82-109.

McEvoy, D. M. (2013). Enforcing compliance with international environmental agreements using a deposit-refund system. International Environmental Agreements: Politics, Law and Economics, 13(4), 481-496.

McEvoy, D. M., Murphy, J. J., Spraggon, J. M., Stranlund, J. K. (2011). The problem of maintaining compliance within stable coalitions: experimental evidence. Oxford Economic Papers, 63(3), 475-498.

McEvoy, D. M., Stranlund, J. K. (2009). Self-enforcing international environmental agreements with costly monitoring for compliance. Environmental Resource and Economics, 42(4), 491–508.

McEvoy, D. M., Stranlund, J. K. (2010). Costly enforcement of voluntary environmental agreements. Environmental Resource and Economics, 47(1), 45-63.

Millner, A., Dietz, S., Heal, G. (2013). Scientific ambiguity and climate policy. Environmental and Resource Economics, 55(1), 21-46.

Montgomery, W. D. (1972). Markets in licenses and efficient pollution control programs. Journal of Economic Theory, 5(3), 395-418.

Morris, J., Paltsev, S., Reilly, J. (2008). Marginal abatement costs and marginal welfare costs for greenhouse gas emissions reductions: results from the EPPA model, MIT Joint Program on the Science and Policy of Global Change, Report 164, Cambridge, MIT.

Muthoo, A. (1999). Bargaining theory with applications. Cambridge and New York: Cambridge University Press.

Nagashima, M., Dellink, R., Van Ierland, E., Weikard, H. P. (2009). Stability of international climate coalitions—a comparison of transfer schemes. Ecological Economics, 68(5), 1476-1487.

Nagashima, M., Weikard, H.-P., de Bruin, K., Dellink, R. (2011). International climate agreements under induced technological change. Metroeconomica, 62 (4), 612-634.

Nash, J.F. (1950). The bargaining problem. Econometrica, 18(2), 155-162.

Nordhaus, W. D. (1994). Managing the Global Commons: The Economics of Climate Change, MIT, Cambridge.

Nordhaus, W. D. (2008). A Question of Balance: Economic Modelling of Global Warming, Yale University Press, New Haven, CT.

Nordhaus, W. D., Yang, Z. (1996). A regional dynamic general-equilibrium model of alternative climate-change strategies. The American Economic Review, 86 (4), 741-765.

Paltsev, S., Reilly, J.M., Jacoby, H.D., Eckaus, R.S., McFarland, J., Sarofim, M., Asadoorian, M., Babiker, M. (2005). The MIT Emissions Prediction and Policy Analysis (EPPA) model: version 4, MIT joint Program on the Science and Policy of Global Change, Report 125, Cambridge, MIT.

Paltsev, S., Morris, J., Cai, Y., Karplus, V., Jacoby, H. (2012). The Role of China in Mitigating Climate Change, MIT Joint Program on the Science and Policy of Global Change, Report 215, Cambridge, MIT.

Polinsky, A. M., Shavell, S. (2000). The economics of public enforcement of law. Journal of Economic Literature, 38(1), 45-76.

Powell, R. (2002). Bargaining theory and international conflict. Annual Review of Political Science, 5, 1-30.

Regan, D. H. (1972). The Problem of Social Cost Revisited. The Journal of Law & Economics, 15(2), 427–437.

Rehdanz, K., Tol, R.S.J. (2005). Unilateral regulation of bilateral trade in greenhouse gas emission permits. Ecological Economics, 54 (4), 397–416.

Rose, A., Stevens, B. (1993). The efficiency and equity of marketable permits for CO_2 emissions. Resource and Energy Economics, 15(1), 117-146.

Rose, A., Stevens, B. (1998). A dynamic analysis of fairness in global warming policies: Kyoto, Buenos Aires, and beyond. Journal of Applied Economics, 1, 329–362.

Rose, A., Stevens, B., Edmonds, J., Wise, M. (1998). International equity and differentiation in global warming policy. Environmental and Resource Economics, 12(1), 25-51.

Rubinstein, A. (1982). Perfect Equilibrium in a Bargaining Model. Econometrica, 50 (1), 97-109.

Sáiz, M. E., Hendrix, E. M., Olieman, N. J. (2006). On the computation of stability in multiple coalition formation games. Computational Economics, 28(3), 251-275.

Schmidt, C. (2000). Designing international environmental agreements. Edward Elgar Publishing, Cheltenham, UK.

Sinclair, D. (1997). Self-regulation versus command and control? Beyond false dichotomies. Law & Policy, 19(4), 529-559.

Stavins, R.N., Jaffe, J. (2008). Linkage of tradable permit systems in international climate policy architecture. Nota di lavoro // Fondazione Eni Enrico Mattei: Sustainable development, No. 90.2008.

Stern, N. (2008). The economics of climate change. American Economic Review, 98(2), 1-37.

Stevens, B., Rose, A. (2002). A dynamic analysis of the marketable permits approach to global warming policy: a comparison of spatial and temporal flexibility. Journal of Environmental Economics and Management, 44, 45-69.

Stranlund, J.K. (2007). The regulatory choice of noncompliance in emissions trading programs. Environmental Resource and Economics, 38, 99–117.

Tietenberg, T.H. (1985). Emissions trading. Resources for the Future, Washington, D.C.

Tol, R. S. J. (1997). A decision-analytic treatise of the enhanced greenhouse effect, dissertation, Vrije Universiteit, Amsterdam.

Tol, R. S. J. (2009). The economic effects of climate change. Journal of Economic Perspectives, 23, 29–51.

UNFCCC (1998). Kyoto Protocol to the United Nations Framework Convention on Climate Change. http://unfccc.int/resource/docs/convkp/kpeng.pdf#page=12.

van der Pol, T., Weikard, H.-P., van Ierland, E. (2012). Can altruism stabilise international climate agreements? Ecological Economics, 81, 112-120.

Wagner, R. H. (1988). Economic interdependence, bargaining power, and political influence. International Organization, 42(3), 461-483.

Wangler, L., Altamirano-Cabrera, J.-C., Weikard, H. P. (2013). The political economy of international environmental agreements: a survey. International Environmental Agreements: Politics, Law and Economics, 13 (3), 387-403.

Weikard, H.-P. (2004). Who should receive the CO2 emission permits? In: Döring, R. and M. Rühs (eds.): Ökonomische Rationalität und praktische Vernunft – Gerechtigkeit, Ökologische Ökonomie und Naturschutz. Festschrift für Ulrich Hampicke. Würzburg: Königshausen und Neumann. 71-82.

Weikard, H.-P. (2009). Cartel stability under an optimal sharing rule. The Manchester School, 77 (5), 575–593.

Weikard, H. P., Dellink, R. (2014). Sticks and carrots for the design of international climate agreements with renegotiations. Annals of Operations Research, 220(1), 49-68.

Weikard, H. P., Finus, M., Altamirano-Cabrera, J. C. (2006). The impact of surplus sharing on the stability of international climate agreements. Oxford Economic Papers, 58(2), 209-232.

Weitzman, M. L. (2009). On modelling and interpreting the economics of catastrophic climate change. The Review of Economics and Statistics, 91(1), 1–19.

Winkelmann, R., Levermann, A., Ridgwell, A., Caldeira, K. (2015). Combustion of available fossil fuel resources sufficient to eliminate the Antarctic Ice Sheet. Science advances, 1(8), e1500589.

Summary

Relying on individual countries' efforts to reduce greenhouse gas (GHG) emissions is insufficient for keeping the rising global average temperature below a level detrimental to humans. Instead, the threat of climate change requires global cooperative efforts to reach significant reductions in GHG emissions. Yet, it is difficult to reach a consensus on sharing the burden of climate change mitigation among nation states because of the free-riding incentives resulting from the public good property of climate change mitigation. In the absence of a supranational agency that takes the role of an enforcer of commitments, the compliance with mitigation commitments is also an obstacle to mitigation cooperation even after an international climate agreement (ICA) has been signed and ratified. To enable effective GHG emissions mitigation, the design of institutional and economic instruments that can facilitate the formation of ICAs is needed. The aim of this thesis is therefore to study the stability and effectiveness of ICAs by considering the impacts of carbon trade, countries' uneven bargaining powers and enforcement mechanisms. Game theoretical modelling is a useful and appropriate tool to study the incentive mechanisms applied to agents with different preferences, and is often used to study the formation and stability of ICAs. Hence, game theory applies throughout this thesis. To relate the analytical solutions to the real problems in ICAs formation and thus to provide policy insights, numerical models and analyses are also employed and presented in this thesis.

While Chapter 1 sets the scene and formulates the research questions, Chapter 2 develops a four-stage coalition formation game in order to study the impact of a carbon market on regional incentives to join an ICA for GHG emissions mitigation when the carbon market is established independently of this agreement. This carbon market, which is assumed to be formed after the mitigation coalition, is open to the mitigation coalition and its outsiders. In particular, the initial emissions permit choice of each participant in the carbon market is based on an endogenous permit choice mechanism introduced by Helm (2003). The impact of this carbon market on incentives to join the mitigation coalition and the interlinkages between the two coalitions are studied. Results show that number and size of stable mitigation coalitions are smaller with than without a carbon market. The intuition is that in the presence of a carbon market, players are faced with more strategic choices implying that members of the stable mitigation coalitions have a potential incentive to deviate from a mitigation coalition and become a carbon trader instead. This can destabilise the mitigation coalition. Because the

carbon market offers an alternative or complementary policy instrument to facilitate mitigation, some non-signatories to the mitigation coalition would join a carbon market which helps to reduce global emissions. However, the mitigation coalition has no incentive to join a carbon market with non-signatories, since the benefits from free riding surpass the net gains from carbon trade. It is concluded that even with this dual-agreement system, free-riding incentives prevail and global cooperation does not emerge.

Chapter 3 develops a model of an international carbon market where allowances are endogenously determined by each member of a carbon trade agreement, but with an exogenous constraint on the number of allowances per member. A global model is used to explore the incentives for regions to participate in such a carbon market and to examine its performance. Results show that the stability and effectiveness of an international carbon market can be improved by imposing constraints on individual allowance choices compared to a carbon market without such constraints. Constraints on allowance choices reduce "hot air" and increase global welfare and mitigation. Under a relatively lax constraint (12% below BAU emissions in the STACO calibration), a carbon market with the largest membership can be formed. When tightening the constraint the stable carbon markets become smaller but perform better in terms of global abatement and welfare. If the constraint is too tight, however, no stable carbon market can be formed. Moreover, numerical results also demonstrate that by tying individual allowance choice constraints to the Nash-emissions levels, international carbon markets are more successful in terms of global welfare and abatement, as it responds better to individual incentives to participate. Hence, a revision of the baseline could ease current climate negotiations, where BAU emissions are still dominant for defining and negotiating abatement targets or emission allowances.

Chapter 4 studies the impact of using the Nash bargaining solution (NBS) for distributing coalitional gains under different sets of bargaining weights on the stability and effectiveness of international climate agreements. International climate negotiations are modelled as a Nash bargaining game in which cooperative gains are distributed based on the NBS with asymmetric bargaining power. In climate negotiations, asymmetric countries' bargaining powers are unequal and may be driven by different characteristics of the players. By discussing and reviewing potential reasons that could induce differences of negotiators' bargaining power in international climate negotiations, five different factors that determine negotiators' bargaining power are identified: i) discount factor; ii) abatement efforts; iii) abatement costs; iv) climate change damages; v) economic power. Numerical results illustrate that the Nash bargaining solution can improve the participation incentives and performances of ICAs as compared to those that do not redistribute gains from cooperation, but its capacity to overcome free-riding incentives is limited. The success of international climate agreements depends on the set of bargaining weights that matters in climate negotiation. Among five sets of asymmetric bargaining weights, the one determined by negotiators' economic power can facilitate a climate coalition that comprises two of the largest emitters (China and United States) jointly with two other regions (High Income Asia countries and Brazil). In climate policy making multiple determinants of bargaining power will play a role. This is because countries' incentives to cooperate on GHG mitigation are impacted in a complex way by factors that are related to abatement options, climate change vulnerability and economic power. Hence, the bargaining power of each negotiator is likely to be driven by multiple determinants. Our model can be extended in a straightforward manner to account for different interlinked drivers of bargaining power if these can be determined.

Chapter 5 studies the design of an optimal enforcement mechanism for a self-enforcing ICA and the impact of an optimally designed enforcement mechanism on participation and compliance. The model is formulated as a sequential cartel formation game, in which the coalition chooses an enforcement policy comprising an abatement target, monitoring expenditures and fines. Individual signatories respond by choosing their preferred abatement level which may or may not comply with the target. In equilibrium, signatories' compliance levels are determined by individual welfare maximization under the agreed enforcement policy. Considering partial compliance, it is demonstrated that the extent of compliance depends on abatement targets, monitoring expenditures (the intensity of monitoring) and the fine. Results show that the compliance level of a coalition member can always be improved by increasing the monitoring expenditure. However, the effect of the target on compliance levels depends on the structure of the fine function. The coalitional best abatement level can be induced as a joint result of the monitoring expenditure and the target when the fine function is quadratic. Because monitoring is costly, full compliance will usually not be enforced. As to the stability, a "broad" and "deep" climate coalition cannot be formed with a costly enforcement mechanism.

Four conclusions can be drawn from this thesis. First, establishing a carbon market outside of a mitigation coalition may not help to improve countries' incentives to join the mitigation coalition, while the global mitigation level can be increased because the carbon market offers an alternative or complementary policy instrument to facilitate mitigation.

Second, imposing a constraint with moderate degrees of strictness on initial allowance choices can increase the mitigation effectiveness and participation level of an international carbon market. Moreover, changing the baseline of the constraint from the BAU emissions level to the non-cooperative Nash level can also be considered a way to improve the success of an international carbon market. Third, factors that could influence bargaining powers of negotiating countries can be brought up and determined at pre-stage of international climate negotiations. Hence, policy designers for negotiations of ICAs may consider to use the determinants of bargaining power that can enhance large emitters' bargaining position to improve the mitigation effectiveness of ICAs. Fourth, due to the existence of free-riding incentives in the process of compliance with mitigation commitments, an optimally designed enforcement policy is needed for the successful implementation of ICAs.

Acknowledgements

When I just started my PhD study, I was stuck with settling on research topics and questions. I was so lucky that Prof. Ekko van Ierland provided me an opportunity to work on my current research topic, which is interesting and attractive to me. Ekko, thanks for your trust, which makes me feel confident and supports me spiritually to finish my PhD. My gratitude also goes to my co-promotors Hans-Peter Weikard and Xueqin Zhu. Hans-Peter, lots of thanks to your guidance in many aspects, from scientific writing, rigorous attitude to scientific research, solid knowledge to cutting-edge research questions. I really appreciate the time and efforts you have been spending on improving my papers and academic abilities. I wish I could become a researcher like you, being precise and dedicated. Xueqin, thank you very much for your encouragement, understanding and help. Thanks for sharing experiences of your PhD life with me and encouraging me when I confronted with difficulties. I can always get positive energy from talking with you. Also thank you for your guidance in programing and modelling by using GAMS. Your hard working attitude keeps inspiring me during my PhD study and also in my future career. I would also like to thank Prof. Michael Finus, who is one of co-authors for one chapter of this thesis. Michael, thank you for giving me a chance to come to Bath University and to work with you and also for your time and efforts invested in the paper. I have been gaining a lot from talking with you about papers and research.

My thanks also go to my colleagues at ENR Group. It was nice to share with you the pains and gains of doing a PhD project. Specially, I would like to say thank you to Wil and Gre, for all your assistance.

I also want to express my gratitude to my Chinese friends. Thank you for making me feel much less lonely and home-sick by eating and hanging out together.

My thanks also go to my husband Thomas and my son Timon. Thomas, thank you for your understanding and support. Your help has made me grow. Thank you Timon, you bring so much joy to our life. I am also grateful for help and support from my parents-in-law, Nel and Leo, and also help from Stefan and Eveline. All of you have given me so much love.

Last but foremost, my deepest love goes to my family: my parents and brothers. Dearest baba and mama, thank you for being open-minded and your unconditional love, otherwise I would not go abroad to pursue and complete my PhD study. I feel so proud and lucky to be your daughter.

About the author

Shumin Yu was born on May 29, 1984 in Zhengzhou, China. She completed her Secondary school at Zhengzhou No.4 Middle School in 2003. In the same year she started her BSc study majored in Electronic business at PLA Information Engineering University in China. After completing BSc study in 2007, she started her MSc in Management Science and Engineering at Northwest Agricultural and Forest University, China. During her MSc studies, she gained expertise in agricultural economics. She obtained the MSc degree in 2011 for her thesis on the development of agricultural informatization in rural China and its impact on agricultural total factor productivity (ATFP). At the same year she received a scholarship from the Erasmus Mundus EURASIA 2 project and began to pursue her PhD studies at Environmental Economics and Natural Resources (ENR) Group of Wageningen University. During PhD trajectory, she had been studying incentive mechanisms in international cooperation on climate change mitigation under the supervision of Prof. Dr Ekko van Ierland, Dr Hans-Peter Weikard and Dr Xueqin Zhu. In 2015, she was awarded the Junior Researcher Grant from the Wageningen School of Social Sciences to conduct a research project on the stability of international climate agreements with costly monitoring and enforcement. This research project is cooperated with Prof. Dr Michael Finus at Bath University in England. The outcome of this research is included as one chapter of this PhD thesis.

Shumin Yu

Wageningen School of Social Sciences (WASS)

Completed Training and Supervision Plan

Name of the learning activity	Department/Institute	Year	ECTS*
A) Project related competences			
Advanced Microeconomics ECH-32306	WUR	2011	6
Advanced Macroeconomics ENR-30806	WUR	2011	6
Economics and Management of Natural Resources ENR-31306	WUR	2012	6
Theories and Models in Environmental Economics ENR-30306	WUR	2012	6
Advanced Econometrics AEP-60306	WUR	2014	6
Modelling critical transitions in nature and society	SENSE	2014	2
B) General research related competences			
Introduction course	WASS	2012	1
Techniques for Writing and Presenting a Scientific Paper	WGS	2012	1.2
'The Effect of Permit Trade on the Stability and Efficiency of International Climate Agreements (ICAs)'	Int. Conf. Cooperation or Conflict? Economics of natural resources and food, Wageningen University, The Netherlands	2013	1
'International climate agreements and the scope for carbon trade with endogenous permits'	5th World Congress of Environmental and Resource Economists, Istanbul, Turkey	2014	1
'International carbon trade with constrained endogenous allowance choices: Results from the STACO model'	21st Annual Conference of the European Association of Environmental and Resource Economists, University of Helsinki, Finland	2015	1
C) Career related competences/personal development			
From Topic to Proposal	WASS	2011	4
Visiting researcher in the University of Bath, UK		2015	2

Total	**43.2**

*One credit according to ECTS is on average equivalent to 28 hours of study load

The research described in this thesis was financially supported by the Erasmus Mundus EURASIA 2 program.