JYVÄSKYLÄ STUDIES IN COMPUTING
94

Yulia Pavlova

Multistage Coalition Formation Game of a Self-Enforcing International Environmental Agreement

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN COMPUTING 94

Yulia Pavlova

Multistage Coalition Formation Game of a Self-Enforcing International Environmental Agreement

Esitetään Jyväskylän yliopiston informaatioteknologian tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Mattilanniemen salissa MaC102
joulukuun 2. päivänä 2008 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Information Technology of the University of Jyväskylä,
in Mattilanniemi, Auditorium MaC102, on December 2, 2008 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2008

Multistage Coalition Formation Game of a Self-Enforcing International Environmental Agreement

JYVÄSKYLÄ STUDIES IN COMPUTING 94

Yulia Pavlova

Multistage Coalition Formation Game of a Self-Enforcing International Environmental Agreement

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2008

Editors
Timo Männikkö
Department of Mathematical Information Technology, University of Jyväskylä
Pekka Olsbo, Marja-Leena Tynkkynen
Publishing Unit, University Library of Jyväskylä

URN:ISBN:978-951-39-3416-3
ISBN 978-951-39-3416-3 (PDF)

ISBN 978-951-39-3400-2 (nid.)
ISSN 1456-5390

Jyväskylä University Printing House, Jyväskylä 2008

ABSTRACT

Pavlova, Yulia
Multistage Coalition Formation Game
of a Self-Enforcing International Environmental Agreement
Jyväskylä: University of Jyväskylä, 2008, 127 p.
(Jyväskylä Studies in Computing
ISSN 1456-5390; 94)
ISBN 978-951-39-3416-3 (PDF), 978-951-39-3400-2 (nid.)
Finnish summary
Diss.

As environmental problems become more and more threatening, political remedies for these problems have become increasingly international. Over the past decade, a variety of multilateral agreements to protect the environment have been initiated, though their effectiveness and continued existence is in peril.

Many of the challenges to these agreements can be explained by the concept of positive externality, meaning that certain countries benefit from the environmentally friendly actions of their neighbors, without making any efforts of their own. Hence, if an agreement places restrictions on a countries technological and economic development, contradicting their self-intrest, it will lead to free-riding in terms of participation and compliance. In order to achieve multilateral cooperation, the strategic possibility interests of the parties must be understood. In this thesis, game theory has been chosen as a tool to analyze these interests and the ways they affect the formation, design, and performance of international environmental agreements (IEAs).

After providing a survey of game-theoretical methodology in the area of IEAs, we examine certain aspects of IEAs such as the choice of emission reduction targets, membership status (signatory versus free-rider), and mechanisms to motivate countries to participate in an agreement. In particular, we address the following question: If the participants and free-riders in an IEA have incentives to change their status during the life-cycle of the agreement, is there a threat to compliance with that agreement? We develop a methodology to achieve the sustainability of an IEA, and prove that under this methodology, participants in the IEA comply with the abatement targets. Furthermore, we asses the impact of the agreement on the welfare of the involved countries and on the global level of pollution.

Keywords: IEA, coalitional game, heterogeneous players, self-enforcement, transfer mechanisms, multistage game, time-consistency, renegotiation.

Author Yulia Pavlova
Department of Mathematical Information Technology
University of Jyväskylä
Finland
jupavlov@jyu.fi

Supervisors Dr. Maria Dementieva
Department of Mathematical Information Technology
University of Jyväskylä
Finland

Professor Pekka Neittaanmäki
Department of Mathematical Information Technology
University of Jyväskylä
Finland

Professor Victor Zakharov
Department of Mathematical Modelling of Energetic Systems
Faculty of Applied Mathematics and Control Processes
St. Petersburg University
Russia

Reviewers Dr. Marko Lindroos
Department of Economics and Management
University of Helsinki
Finland

Professor Georges Zaccour
Department of Marketing
HEC Montréal
Canada

Opponent Professor Rabah Amir
Department of Economics
University of Arizona
United States

ACKNOWLEDGEMENTS

These words of gratitude are devoted to all those people who helped to make this thesis possible.

First of all, I would like to thank my supervisors, Prof. Pekka Neittaanmäki, Prof. Victor Zakharov and Dr. Maria Dementieva for their help and support. It was Victor Zakharov who inspired me with an idea to apply game-theoretic modeling to analyze environmental problems. I am in debt to Victor Zakharov and Maria Dementieva for their scientific guidance and encouragement. I benefitted from our numerous discussions and received many comments that substantially improved the choice of the topics and the style of the resulting presentation. I am grateful to Pekka Neitaanmäki for his trust and kind approval. He encouraged and provided me with an ideal working environment, so that I could conduct my research and write this manuscript. I would also like to thank my colleagues Tero Tuovinen, Timo Aittokoski, Iryna Yevseyeva and Elena Gorshkova for their help and friendship. I also wish to acknowledge funding from the COMAS Graduate School of the University of Jyväskylä, which made this thesis possible.

My special thanks go also to the reviewers of my doctoral thesis, Prof. Georges Zaccour and Dr. Marko Lindroos, for their careful reading, inspiration, and positive references. Thanks to Prof. Rabah Amir, who kindly agreed to be my opponent at the defense of this thesis. I also wish to thank Georges Zaccour for his suggestions on taking a global understanding of my research results and Marko Lindroos for his extensive remarks on the various stages of preparation for this manuscript.

I have also benefitted from discussions with Michèle Breton, Arkady Kryazhymskiy, Fabian Wagner, Brian Fath, Raimo Hämäläinen, Michael Finus, Leon Petrosjan, Jürgen Sheffran, Vladimir Mazalov, Anatoly Zenkevich, Rabah Amir and from the remarks of participants of the 12th and 13th ISDG, EURO XXII, GTM 2007 and 2008, 7th Meeting on Game Theory and Practice Dedicated to Energy, Environment and Natural Resources.

I am in debt to my family, my mother and father, Larisa and Nikolay Pavlov, for their understanding and endless support during these years that I have pursued my doctorate. At times, it has been good to know they have been just a phone call or an email away. I am especially grateful to my father who, being a scientist himself, gave me an outstanding example of living wisdom and dedication to the research profession.

Finally, I wish to thank Tapio Kymäläinen: All I can say is that it would take another thesis to express my gratitude for him. His patience, love and encouragement have upheld me, particularly during those many days in which I spent more time with my computer than with Tapio.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

The increasing urgency of environmental problems has attracted the attention of scientists, politicians, and society at large. Environmental issues include the depletion of the ozone layer, the loss of biodiversity, and the affects of climate change. To protect the environment and insure the stability of the ecosystem, a variety of international environmental agreements (IEAs) have been developed. These documents prescribe, among other things, pollution limits, the growth of industrial efficiency, and the use of non-hazardous and environmentally friendly raw materials and fuels. They also recommend more careful usage of resources and the restoration of a balanced ecosystem.

Though the parties to such agreements are often well-intentioned, IEA's negotiation and enforcement face significant obstacles, and they often fail to achieve the desired results. At the root of such problems lies the realistic principle that countries act only in their own interest. It is obvious that the required shifts in industrial technology are costly in an economic and social sense, and thus each country wants to avoid paying for environmental protection, even while recognizing that if every country does so, the overall result is not satisfactory.

Despite the fact that rational debate on emission reduction can lead to an agreement prescribing each country to reduce emission, self-interest and the sovereignty of each country would put such an agreement in jeopardy. This can occur in several ways. First, it is possible that an agreement is accessed and complied with by all potential parties, but the outcome is nevertheless insufficient. Second, the agreement might be accessed by only a few nations, or the parties involved do not honor their obligations. Obstacles of this second kind are referred to as free-riding.

In order to be successful, the agreement must be *self-enforcing* (or *self-enforceable*, [Brechet *et al.* 2007]). This concept entails three properties, which are also present in most stability criteria [Barrett 2003]. *Individual rationality* means that no agreement member can gain by withdrawing from the agreement, given the choices of the other members. Moreover, no non-member can gain by joining the agreement, again, given the choices of the other countries. This property also applies to compliance, *i.e.*, no member can be better off by failing to comply with its commitment, given the agreements design, and no non-member can gain by changing (positively or negatively) its environmental practices. *Collective rationality* means that the joint decisions made by the agreement members are such that their total gain would not increase by collectively deviating from the chosen targets. Finally, *fairness* means that the agreement should set feasible goals and obligations without prejudice against any single country. However, this property is of minor importance in comparison with the first two, as an agreement which does not have the aforementioned properties has no chance of success.

Although collective rationality can call for global cooperation, individual rationality often stipulates free-riding behavior with respect to participation and/or compliance. To mitigate this free-riding, the strategic options and incentives of the countries must be understood. Game theory is the perfect instrument for accomplishing this task.

In game theory, “games" are used to model serious human interactions, such as market competition, arms races, and pollution emissions. In these interactions, as in a game, the individual's choices constitute a strategy, and the outcome depends on the strategy chosen by each participant. In this light, game theory allows for a complex study of the formation and stability of IEAs.

In this thesis, an IEA is interpreted as a coalitional game, the coalition of players represents the member-countries of the agreement. The central issues discussed here correspond to the formation and promotion of international cooperation among a heterogeneous group of nations towards the goal of pollution control. We also examine the mechanisms enhancing such collaboration, such as side payments and emission trading. Our results concern the dynamics of the agreement, and in particular, the following question: if both members and free-riders have incentives to change their status during the agreement life-cycle, is there a threat to compliance with the agreement?

In Chapter 2, we introduce some basic terms and notions, and give a survey of some well-known concepts in game theory. We also present contemporary results in a critical way, build bridges between various results, and outline some promising directions in the area of environmental agreements.

Part I describes our original contributions to the area of static games. These results address IEA formation and the redistribution of gains from cooperation among signatories. We explore an economic-ecological model of an $\mathcal{N}$-country world, linking the economic activity of the countries with the physical state of the environment. This link is given by the social welfare function, which is the difference between the profit from emission reduction and the environmental protection costs. A detailed description of the basic assumptions and model characteristics can be found in Section 3.1. Following this approach, we examine a game of heterogeneous players, assuming *ex ante* that they are collected in several groups regarding their welfare function.

Section 3.2 explores a particular case of agreement formation, where the players are split into two groups. We consider the extreme cases of full cooperation (a grand coalition forms) and pure non-cooperation (all players act alone), and compare the environmental and economic benefits, arguing the preferability of full cooperation. As we have already mentioned, full cooperation is rarely achievable due to individual rationality, which motivates players to withdraw from the agreement and benefit from the collective efforts of others.

Such reasoning leads to the consideration of an agreement with partial participation of the players, and to the analysis of the coalition formations process, which divides players into signatories and free-riders. We determine the optimal abatement levels of the players and characterize the structure of a stable coalition by applying the principle of *self-enforcement*. This provides an insight into the expected environmental benefits of such an agreement and the players' welfare.

The main result of Chapter 3 is presented in Section 3.3, where the approach initiated in Section 3.2 is extended to the heterogeneous case, in which the players are split into K types. Section 3.4 provides an analysis of the relationship between the size of the agreement and the achieved emission reduction.

The heterogeneity of the players gives an opportunity to launch transfer scheme mechanisms, which ought to enhance the players' commitment to the IEA and reduce the incentives to free-ride. This analysis is presented in Chapter 4.

Part II is devoted to the dynamic performance of the chosen abatement commitments and dynamic stability of the agremeent. We focus on agreement dynamics, and, in particular, the following intuitive question: If the participants and free-riders in an IEA have incentives to change their status during the life-cycle of the agreement, is there a threat to compliance with the agreement? Furthermore, we ask if the dynamics of pollution flow associated with the abatement activities can affect agreement stability. Due to the present dynamics, the property of time-consistency [Petrosjan 1977] is of great importance and plays fundamental role in the undertaken analysis.

The dynamics of pollution flow is related to the way in which the countries decide to reduce emissions during the life-cycle of the agreement. In order to realistically model countries' behavior, we consider the following approach to stepwise emission reduction. The emission reduction over each stage is chosen to maximize each country's welfare both over the current stage and over the rest of the abatement path. We prove that such a scheme is time-consistent. Moreover, it appears to yield internal dynamic stability, and so guarantees that the signatories have no incentives to leave the agreement, given that no new member can access it. Part II concludes with an estimate of the players' welfare over the accounted period.

The issue of agreement vulnerability is continued in Part II. The analysis of an IEA's dynamic external stability presented in Section 6.3 reveals that with the multistage abatement scheme described above, external dynamic stability is violated after a certain time threshold. This provides countries with incentives to withdraw from or access the agreement. Repeated IEA negotiation is needed to handle this potential vulnerability, including the reconsideration of each countries' membership status and emission reduction targets. Our analysis shows that renegotiation eventually reassigns the abatement commitments to the agreement members, but leaves the agreement structure unaffected. Moreover, it appears that renegotiations take place regularly, sufficiently increasing the total emissions reduction.

Finally, we present concluding remarks and introduce some potential directions for further research. These include the analysis of the dynamic stability of the agreement upon meeting the agreement targets, and the potential internal stability of the agreement in light of the dynamic framework.

2 A SURVEY

The game theory gives mathematical representations of agents' preferences and makes predictions about agents' outcomes, resulting from their interaction. There are two self-sustained directions of game-theoretic literature[1] that explore the formation mechanism of IEAs.

One stream of works utilizes a cooperative approach. The traditional cooperative approach is based on the fundamental assumption that players have already agreed to cooperate and the objective is to maximize the joint payoff of the game. It implies existence of a 'third party' (a supernational authority, which has power to bring an IEA into force).

Another way of describing IEA belongs to non-cooperative game theory, where players behave according to rational self-interest and cannot communicate before the game. This approach allows a group of players to cooperate and form a coalition which can be smaller than the grand coalition (all players).

Both cooperative and non-cooperative models test whether coalition members have incentives to leave the coalition, considering reaction of remaining players that imply a punishment. As well as cooperative models, in most non-cooperative models the choice of abatement strategies is made by players under assumption of potential collaboration. Another similarity lies in the facts that both approaches expect coalition members, first, to cooperate among themselves and to act competitively (non-cooperatively) against others, and second, to fully comply with agreement obligations once a coalition has formed.

For cooperative and coalition games, stability has to be evaluated after the solution to the game (abatement strategies) is determined. Cooperative game theory is associated with the stability concept of the core and its variations. The joint payoffs of the players are given with help of a corresponding characteristic function. It is shown that an IEA signed by all countries is stable, using the γ-core concept and implementing transfers to solve heterogeneity of the countries involved, [Chandler & Tulkens 1995] and [Chandler & Tulkens 1997]. Usually non-cooperative game theory is associated with the concept of

[1] A review of current literature see [Finus 2001], [Petrosjan & Zakharov], [Carraro & Siniscalco 1998], [Ioannidis 2000], [Carraro *et al.* 2006].

internal and external stability and its modifications, [Hoel 1992], [Bauer 1992], [Carraro & Siniscalco 1993], [Barrett 1994a]. The idea is to check for which size of the coalition an individual country is indifferent, between joining or leaving and to use that as the stability concept. It reveals that the grand coalition will not generally be stable as individual players have a strong incentive to take a free-rider position, and the size of a stable coalition is typically very small[2].

The explanation for such opposing conclusions is the difference in behavioral assumptions. For instance, the γ-core concept assumes that in case a sub-coalition deviates, the grand coalition falls apart and the rest of the players act non-cooperatively against the sub-coalition. The non-cooperative approach supposes that only a single player can deviate from a coalition (this assumption can be considered as a drawback), but on the other hand if it happens the other members continue as a coalition (more realistic outcome), which yields sufficient free-rider benefits to end up with only a small coalition.

The present work utilizes the non-cooperative approach due to the following reasoning. First of all, the assumption of an absence of a supernational authority seems more realistic. It allows reference to questions of formation of a successful agreement starting from individual incentives, but not from collective rationality, which looks natural while dealing with problems of international environmental cooperation.

Second, coalitional games of environmental protection are games with *positive externality*. Technically it means that the merging of a larger coalition from the smaller ones and/or single players creates a positive side effect on those actors who were not involved. This is quite obvious since a cleaner environment certainly benefits society as a whole, but does not increase profits for the party responsible for it. Given the presence of positive externality, cooperative game theory can become less satisfactory to work within, since it suggests that the coalition typically acts independently of the actions chosen by non-members[3], [Yi 2003].

Besides this classification, there are also two possible aspects of free-rider incentives, [Finus 2003a]. The first type refers to the incentive not to participate in an IEA; the second type to the incentive not to comply with the obligations agreed upon an IEA. *Membership models* focus on the first type of free-rider incentive. Implicitly, they assume that once a country joins an agreement it will comply with the agreed treaty obligations. Thus, compliance is exogenous. In contrast, *compliance models* concentrate on the second type of free-rider incentive, starting from the exogenous assumption that some coalition has formed and test whether treaty obligations can be enforced with credible threats to sanction non-compliance. Normally, for modelling an IEA, one or another type of free-riding is accepted. One of the preliminary attempts to combine both free-riding aspects can be found in [de Zeeuw 2008]. In the present work, our analysis mainly takes place in the frame of membership model, though it will be shown in Part II that

2 More detailed analysis of literature is presented in Subsection 2.1.

3 Though it is necessary to mention that these days there are new sharing rules appearing in cooperative framework, which take into account presence of externality effect, for instance, *satisfactory nucleolus* proposed in [Kronbak & Lindroos] in the fishery context.

issue of emission reduction compliance comes to light in the context of IEA dynamics.

2.1 Non-cooperative Membership Models

In non-cooperative membership models countries play a two stage game. Here a term "stage" is used to indicate sequence of players decisions within a static framework. In the first stage, each country decides to join or not to join an IEA. In the second stage, every country decides on emissions (or abatements). The problem of coalition formation is solved backwards and, within each stage, different assumptions (illustrated in Table 2.1) can be made.

TABLE 1 Structure of Coalition Formation in Membership Models

1. Stage: participation

sequence	simultaneous	sequential no revision on members *vs.* revision of members
agreements	single	multiple
membership	open	exclusive majority *vs.* unanimity

2. Stage: abatement and transfers

sequence	simultaneous (Cournot)	sequential (Stackelberg)
abatement	joint welfare maximization (efficient)	bargaining majority *vs.* unanimity
transfers	no	yes
payoffs	objective social planner material certain	subjective political non-material uncertain

A keystone idea, underlying in our analysis, is based on *a conjectural variation model*, ([d'Aspremont *et al.* 1983], [d'Aspremont *et al.* 1986]). This type of model was first used in the context of the IEAs formation in [Carraro & Siniscalco 1993], [Barrett 1991], [Barrett 1992]. The common features of the conjectural variation model are as follows:

(a) they consider only possible deviation from an equilibrium coalition by a single country and discard the possibility that a sub-group of countries may deviate;

(b) this model considers only one coalition (thus all players are divided into two groups: signatories and free-riders) and all non-members behave as

singletons;

(c) equilibrium strategies are based on the myopic behavior of players, who consider only the immediate reaction of fellow players to a change in their strategy but not the subsequent moves (chain reactions) which may be trigged;

(d) choice of coalition abatement target is based on maximization of aggregate welfare function, and abatement decision of free-riders is made by maximizing individual welfare;

(e) besides that, players' payoffs are material and certain.

Alternatively, these games are defined as *open-membership games*. Typical set up for these models is a *two stage game*. In the first stage players decide whether to participate in an agreement or not. It is assumed that this is a binary choice: 'join' and 'do not join'. In the second stage players choose their emission reduction level. The problem is solved backwards.

Sequence of abatement moves

In the literature there are basically two assumptions regarding the sequence of moves: (a) players choose their strategies simultaneously, [Hoel 1992], [Carraro & Siniscalco 1993], [Bauer 1992]; (b) players choose their participation strategy in stage one simultaneously, but abatement levels in the second stage are decided sequentially [Barrett 1994a], [Barrett 1991], [Barrett 1992], [Barrett 1997a]. The first assumption is referred to as *Nash-Cournot assumption* and the second one as the *Stackelberg assumption*. The latter assumption is only in use for the case, when there is a single IEA and the rest of players are free-riders, and implies, that there is a Stackelberg leader (a coalition of signatories), who takes into account the optimal choice of non-signatories that behave as Stackelberg followers. Participants have an advantage towards non-participants as they choose their emission levels based on reaction function of non-signatories. In context of international environmental cooperation, such setting can be justified by arguing that signatories are better informed than non-signatories about emission levels of other countries since they coordinate their environmental policies within an IEA.

In the aforementioned literature, abatement targets are chosen by maximizing global welfare, assuming players to be homogeneous and their payoffs to be material and certain. It is shown that regardless of sequence of moves only coalitions of small size are stable, if gains from cooperation are large. In particular, stable coalition size is $2, 3, 4$ players in [Barrett 1994a], [Diamantoudi & Sartzetakis 2006] and 2 and 3 players in [Carraro & Siniscalco 1993].

Coalition stability

In the Introduction we have requirements, which an agreement must satisfy for self-enforcing: individual rationality, collective rationality and fairness. In a context of a *conjectural variation coalition equilibrium* these requirements are expressed with the help of the following conditions:

(a) *internal stability* (there is no incentive for a signatory to leave the coalition);

(b) *external stability* (there is no incentive for a non-signatory to join the coalition);

(c) *profitability of a coalition* (each signatory is better off than in case of no agreement).

Briefly speaking, the idea of this stability concept is to check for which structure (size and types of players, if they are heterogeneous) of the agreement an individual country is indifferent, between joining or leaving. As we have pointed out above, employing this type of stability leads typically to a coalition of small size.

Transfers

For symmetric (homogeneous) countries the choice of abatement level is based on welfare maximization, and the welfare allocation is trivial: each county receives the same payoff and no reallocation of payoffs is necessary. With heterogeneous players in particular, when assuming countries differ in abatement cost, a coalition can exploit cheap abatement options if a low cost player joins the coalition (thus, one may conclude that low cost countries are attractive as coalition partners).

To create incentives for low cost countries to join the coalition, the sharing of the coalition payoff among members can be used. First analytical and empirical attempts to address this issue are presented in [Hoel 1992], [Carraro & Siniscalco 1993], [Bauer 1992], [Botteon & Carraro 1997], [Barrett 1997a]. For example, [Carraro & Siniscalco 1993], [Barrett 1997a] and [Botteon & Carraro 1997] show that the size of the coalition can be extended if the signatories offer a transfer to outsiders for their willingness to join. More recent studies have addressed the impacts of different sharing rules on the stability of international environmental agreements.

It was shown that if sharing rules are applied to abatement costs, [Bosello *et al.* 2003], or tradable pollution permits, *e.g.* [Altamirano-Cabrera & Finus 2006], there is no guarantee that payoffs satisfy the individual rationality constraint. By contrast, if sharing is applied to the gains from cooperation, [Weikard *et al.* 2006], individual rationality is always satisfied as long as a coalition is profitable. Recently a class of sharing rules has been proposed that divides the difference between the coalition payoff and the sum of the outside option' payoffs[4] of coalition members, [Carraro *et al.* 2006], [Eyckmans & Finus 2004], [Weikard 2005].

[4] The 'outside option' payoff is the payoff a player would receive when leaving the agreement.

In addition to leadership behavior (Stackelberg concept) and transfers, it is necessary to mention that including reputation effects, [Jeppensen & Andersen 1998], issue linkages, [Carraro & Siniscalco 1998], [Barrett 1997b], [Botteon & Carraro 1998], [Le Breton & Soubeyran 1997], [Carraro & Marchiory 2004], [Katsoulacos 1997], [Finus & Rundshagen 2000], setting a minimum participation clause, [Carraro *et al.* 2003], or low emission targets, [Finus 2004], also increase incentives to cooperate.

2.2 Criticism and Alternative Concepts

Further we are going to present some alternatives and extensions to the concepts introduced in the previous section.

2.2.1 Myopic and farsighted behavior

Recently the concept of *farsightedness* was introduced and applied to the problem of IEA's stability, [Diamantoudi & Sartzetakis 2006], [Chwe 1994], [Ray &Vohra 1999], [Eyckmans 2003]. The behavioral assumptions in this model reconcile the cooperative and non-cooperative approaches (described in Chapter 2). Neither is it assumed that the coalition fully breaks down nor that the remaining coalition stays intact but it is assumed that if a country leaves, it may also trigger other countries to leave until some new stable situation is reached.

A country has to compare its initial position with its position at the end of the process and this makes both large and small coalitions possible. At the same time, necessity of tracing a series of deviations can make a game with heterogeneous players rather complex.

2.2.2 Membership: open *vs.* exclusive

In contrast to open membership games (when external players can freely access an agreement if s/he has incentives to do so), games with *exclusive membership* imply that if an outsider to a coalition wants to join, this can be turned down if a majority of players (*majority* voting) or one member (*unanimity* voting) is against.

2.2.3 Agreements: single *vs.* poly

The alternative approach encourage to explore formation of multiple agreements instead of a single agreement. The idea comes from the following reasoning: if the target of getting as many countries as possible into one agreement seems difficult to achieve in the presence of free-rider incentives, then it is practical to allow for several separate agreements among regions of similar interests to foster the success of international cooperation. The proposed game-theoretical framework suggests that once a particular poly-coalition structure has formed (in the first stage), coalition members cooperate among themselves and play non-cooperative against other coalitions (in the second stage).

For analyzing poly-coalition formation process (coalitions are disjoint[5].), a *per-membership partition function* (see [Bloch 1997], [Yi 1997]) is traditionally applied. By definition, this function is a mapping, which assigns a vector of players' individual payoffs to each poly-coalition structure. Presented in [Bloch 1997] and [Yi 1997] characterization of the per-membership partition function for the games with *positive externality* is composed of 4 inequalities, which set conditions on players' payoffs.

The idea to consider multiple coalitions in the context of IEAs goes back to [Carraro 2000]. Recently this approach has mainly been explored for *ex ante* homogeneous countries and an equal sharing scheme. An analytical analysis of single versus multiple coalitions under various membership rules can be found in [Carraro & Marchiory 2003], [Finus 2003b], [Finus & Rundshagen 2003].

Multiple coalition models can be structured in the following already familiar manner. First, membership can be formed by sequential and simultaneous accession. The type of model dictates choice of stability concept. If we deal with *simultaneous move* membership game, such concepts as

- the *strong Nash equilibrium*[6], [Aumann 1959] and
- the *coalition-proof Nash equilibrium*[7] [Bernheim *et el.* 1987]

are applied (as well as some forms of *core-stability* for the cooperative case). Second, membership can be opened or exclusive (*unanimity* or *majority veto*). The most promising among simultaneous move concepts appears to be one based on principle of *farsightedness*, discussed above. The *farsighted coalition stability*, implying *de facto* exclusive membership, delivers Pareto-dominant poly-coalition structure, allowing for sequential deviation and best-reply strategy for external players.

Examples of *sequential move* membership games are

- the *equilibrium binding agreement*s, [Ray & Vohra 1997], and
- the *sequential coalition formations*, introduced in [Bloch 1995], [Bloch 1996] and its developments. For example, in [Ray &Vohra 1999] the fixed sharing rule was improved by making it endogenous, effect of issue linkage was studied in [Finus & Rundshagen 2000] and a role of international coordinator was analyzed in [Finus & Rundshagen 2006].

---

5 The latest efforts are also directed to consider overlapping or intersecting coalitions, [Breton *et al.* 2008a], [Alcalde & Revilla 2001], [Le Breton *et al.* 2007]

6 No subgroup of players can increase its benefit by deviating.

7 The stability concept is similar to the *strong Nash equilibrium* but it rules out less preferable Pareto-dominated equilibrium outcomes which can appear in the first case.

The latter type of game is also called *sequential move unanimity games*, or *SMUG*. In comparison to simultaneous move games, *SMUG* explicitly depicts how coalition is formed and how players coordinate on an equilibrium. In most of the games there is an initiator, who is interested in starting an agreement and proposes a certain coalition. External players are invited to join the coalition only if all the coalition members unanimously agree.

Analysis of the presented approach allows one to conclude (*e.g.* [Finus 2001], [Finus 2007]) that in the games with poly-coalition, more players (than in single coalition models) are involved into bi- or multi-literal collaboration, which results in bigger environmental benefit. Besides that, it is shown that the grand coalition becomes stable in some cases. Comparison of open and exclusive types of membership is in favor of the latter, since it produces larger coalitions (by preventing undesirable deviations). Considering formation of multiple coalitions among heterogeneous players is a potential extension of the current results.

With no distraction from the merits of poly-coalition approach, we would continue our work and analysis of games with a single coalition and open membership: first, all current IEAs on pollution reduction and CO_2 mitigation are single agreements (for instance, there is only one Montreal and one Kyoto protocol), second, it is common that environmental agreements do not restrict accession and hence follow the principle of open membership.

2.2.4 Dynamic Models

The majority of non-cooperative game theory literature uses a static framework to analyze the formation process of a self-enforcing agreement and choice of abatement targets (the detailed discussion of this issue is presented in Sections 2.1 and 3.3.3). When we turn to the question how to fulfill abatement commitments, which were assigned to or adopted by the players, a dynamic model is necessary. Such a framework could not only allow us to consider emission reduction flow but also to trace changes in players' preferences of their agreement status. The former aspect recalls that transboundary environmental damage is usually related to accumulation of pollution, rather than to emission itself. The latter aspect is that countries can revise their decisions of being in or out of an agreement at different points in time if environmental damage (benefit) and cost are changing over time.

Three types of dynamic games have been used to analyze the stability of an IEA in a dynamic framework: the repeated games, see for instance [Barrett 1994a], [Finus & Rundshagen 1998] – [Asheim *et al.* 2006], the differential (difference) games, [Germain *et al.* 2003] – [Rubio & Ulph 2007], and the multistage games, [Zakharov 1988], [Dementieva 2004]. The latter approach has been utilized to consider dynamics and evolution of an IEA in Part II of the present work.

The earliest attempts to consider dynamic models focus either only on the stock of accumulated pollution, assuming IEA membership to be *fixed*, [Rubio & Casino], [Eyckmans 2001], or on coalition membership dynamics (sequential formation process according to the sequential move unanimity game, see Section 2.2.3). Recently attention to the dynamic difference games of agreement formation has substantially increased and the following problems have been studied:

(a) determining the stable pollution stock and the correspondent size stable coalition, [Rubio & Ulph 2007], [Breton *et al.* 2008b],

(b) identifying conditions leading to breaking apart of the agreement, [de Zeeuw 2008],

(c) introducing mechanisms of "sticks and carrots", which could prevent it, [Weikerd & Dellink 2008] .

On the other hand such questions as

(a) specifying a time-consistent scheme of optimal pollution reduction and correspondent dynamics of pollution flow,

(b) analyzing free-riding aspects,

(c) verification of time-consistency of a stable agreement,

(d) intermediate renegotiation of abatement commitments, specified by the agreement, and players' membership status,

have not been sufficiently explored. These and other related topics would be explored in Parts II and III of the present work.

Part I

Static Game

3 MODEL OF IEA

Before going over to issues of dynamic performance of an international environmental agreement, we start in the area of static games, IEAs formation and the redistribution of gains from cooperation among signatories, and describe our original contributions in light of the past results.

The following analysis of agreement formation is based on a non-cooperative approach, implementing a *conjectural variation model* of membership, [Hoel 1992], [Carraro & Siniscalco 1993], [Bauer 1992], [Barrett 1994a], [d'Aspremont *et al.* 1983], [d'Aspremont *et al.* 1986]. Typically it is set up as a two-stage game, where a term "stage" is used to indicate sequence of players' decisions within a static framework. In the first stage, each country decides to join or not to join an IEA. In the second stage, every country decides on abatements. Material of Part I is based on publications [Demetieva & Pavlova 2007a], [Dementieva & Pavlova 2007b], [Pavlova *et al.* 2008].

According to model characterization provided in Table 2.1, and main and alternative assumptions related to membership models discussed in the Chapter 2, we specify our approach in the following manner:

- coalition formation process is restricted to signatories, which coordinate their strategies, and free-riders, behaving as singletons, [Carraro & Siniscalco 1998];
- when withdrawing or accessing the coalition, a player assumes that all other players maintain their status, that allows only singleton movements as it is employed in the concept of internal/external stability (*i.e.* self-enforcement), given in [d'Aspremont *et al.* 1983];
- players make their participation decisions myopically, *i.e.* don't foresee the subsequent chain reaction by other players;
- membership decisions are simultaneously taken in the first stage, and choice of abatement strategy is sequentially made in the second stage;

- within the coalition, players play cooperatively while the coalition and single countries compete in a non cooperative way among them;

- welfare function is objective, material and certain. This assumption is mainly caused by other model characteristics, because the chosen stability concept had to rely on specific payoff functions to judge about signatory number and type; moreover, we suppose that players' payoffs are presented as the difference between polynomial benefit function and quadratic cost function and that all players are familiar with benefits of others;

- heterogeneous players are allocated K different groups so that they can be identical within each group.

Heterogeneous approach, performed by numerical simulations with two groups of players, has been presented in [Barrett 2001]. In [McGinty 2006] the N asymmetric player game is presented. Such an advanced approach allows to study pollution transfers and rule of fair surplus allocation among IEA signatories, though it makes it rather difficult to provide estimations of IEA size and structure. To derive a pattern of a stable IEA, we direct our attention to the game of heterogeneous players (they are allocated among several groups, regarding their welfare function) and generalize principle of *self-enforcement*. In Sections 3.2 and 3.3 we determine optimal abatement levels, characterize structure of stable coalition and get an insight into expected environmental benefits and players' welfare. If players are split into two groups, like in Section 3.2, it may be interpreted as belonging to Annex B countries (Australia, Austria, Belgium, Bulgaria, Canada, Croatia, Czech Republic, Denmark, Estonia, Finland, France (including Monaco), Germany, Greece, Hungary, Iceland, Ireland, Italy (including San Marino), Japan, Latvia, Lithuania, Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russian Federation, Slovakia, Slovenia, Spain, Sweden, Switzerland (including Liechtenstein), Ukraine, United Kingdom, United States of America) and non-Annex countries of the Kyoto Protocol. If we consider three different groups (see Section 3.3 where K types are considered) then it means that we distinguish the following groups: industrialized countries, like USA, European Union, Japan, without pollution permits; rapidly developing countries, like Russia, China and India, whose pollution permits stock is big enough but emission trading might be inefficient because pollution permits would be necessary for internal use to compensate extra emission discharge caused by industrial growth; and agricultural countries with low abatement cost and large pollution permits stock.

Furthermore, heterogeneity of players gives opportunity to launch transfer scheme mechanisms, which should enhance players' commitment to the IEA and reduce free-riding incentives. Such analysis is presented in Chapter 4.

A mechanism of side payments allows one to reshape the agreement structure and attract nations, which so far preferred to be outsiders. We suggest a rule to share a surplus gained by all coalition members, which would guarantee that each nation receives at least as much as it would get deviating from IEA plus a coalition surplus share. We suppose that under such an approach nations prefer to access IEA than to free-ride and IEAs become capable to provide more significant and meaningful impact on environment. Since more nations are involved into the agreement, welfare of elder signatories might drop down, however being still high enough to keep coalition membership profitable. To avoid likely thwarting, we imitate a special committee to be responsible for assigning abatement commitments and establishing pollution permits for coalition members, thereby allowing thus emission trading to start among nations, [Hanley *et al.* 1997]. The following Section 3.1 precisely describes model features.

3.1 Linear Marginal Abatement Benefits and Costs

Let $\mathcal{N} = \bigcup_{i=1}^{K} \mathcal{N}_i \quad (\mathcal{N}_i \cap \mathcal{N}_j = \varnothing, i \neq j)$ be a set of heterogeneous players, *e.g.* countries of the world, each of which emits pollutants that damages a shared environmental resource. Each subset $\mathcal{N}_i, i = 1, \ldots, K$, consists of N_i players of type i, which have similar payoff functions. Thus, the set $\mathcal{N}$ is composed of N elements (players), where $N = \sum_{i=1}^{K} N_i$.

Let set S ($\varnothing \neq S \subseteq \mathcal{N}$) be a coalition of players that jointly intend to reduce their emissions. Players simultaneously and voluntarily decide to join the coalition S or act independently. Denote n_i ($n_i \leq N_i, \quad i = 1, \ldots, K$) as the number of players of type i that joined the coalition. The vector $\mathbf{n} = (n_1, \ldots, n_i, \ldots, n_K)$ describes the IEA structure S. Let us introduce the following practical notations:

- $F = \mathcal{N} \setminus S$ is the set of free-riders (players, who did not join the agreement);
- q_i^S and q_i^F are the individual abatement commitments chosen by a player of type i from the coalition S and a free-rider of type i from the set F, respectively;
- $\mathbf{q}^S = (q_1^S, \ldots, q_K^S)$ and $\mathbf{q}^F = (q_1^F, \ldots, q_K^F)$;
- $Q_S = \sum_{i=1}^{K} n_i q_i^S$ and $Q_F = \sum_{i=1}^{K} (N_i - n_i) q_i^F$ are the abatements that all signatories of the coalition S and all free-riders from the set F commit to reduce;
- $Q = Q_S + Q_F$ is total abatement by all players upon the IEA.

We declare a link between the economic activity of the countries and physical state of environment. Such a link is established through a social welfare (or payoff) function and expressed via an economical-ecological model of the world. The net benefit $\pi_i(\mathbf{q}^S, \mathbf{q}^F)$ of each player of type i, $i = 1, \ldots, K$, depends on its own abatement commitments q_i^S and q_i^F and on emission reduction Q undertaken by all players, [Barrett 1994a],

$$\pi_i(\mathbf{q}^S, \mathbf{q}^F) = B(Q) - C_i(q_i), \tag{1}$$

and be expressed as difference between benefit and cost functions. The current abatement benefits $B(Q)$ are assumed to be identical for all participants and depend on the current total abatement Q as follows

$$B(Q) = \frac{b}{N}(aQ - Q^2/2). \tag{2}$$

The positive parameters a and b describe the current pollution and slope of the marginal benefit function. We assume that environmental benefit (2) is equally allocated among all countries.

According to (2) the individual marginal benefit function (which can be found as the differential of $B(Q)$) is

$$MB(Q) = \frac{b}{N}(a - Q)$$

and the global marginal benefit function (which can be found as the differential of global benefit function) is

$$MB(Q)N = b(a - Q).$$

Each country's abatement costs depend on its own abatement level q_i, $i = 1, \ldots, K$. For the country of type i the abatement cost function $C_i(q_i)$ is assumed to be given by

$$C_i(q_i) = \frac{1}{2}c_i q_j^2, \tag{3}$$

where q_i is each country's abatement and parameter $c_i > 0$ equals the slope of each country's marginal abatement cost curve. The marginal cost function is

$$MC_i(q_i) = c_i q_i.$$

We interpret agreement formation as a static two-level game $\Gamma_0(S) = \langle \mathcal{N}, \{q_i^S, q_i^F\}_{i=1}^K, \{\pi_i^S, \pi_i^F\}_{i=1}^K \rangle$, where q_i^S and q_i^F are the players' strategies and π_i^S and π_i^F are the net benefits of players of type i. In the current game$\Gamma_0(S)$ the coalition S is the leader, and the free-riders in F are the followers. The strategies (abatement targets) $(\mathbf{q}^S, \mathbf{q}^F)$ are said to be feasible if $Q \leq a$.

3.2 Coalition Formation Game Among Players of Two Types

It is quite obvious that considering formation of an IEA as a non-cooperative game and potential countries-signatories as players, we may run against a problem which is too complicated. Indeed, each particular nation requires unique adjustment of benefit and cost function parameters. This in turn allows to estimate, quite explicitly, individual abatements and the net benefits of each country, but when the question concerns stability of agreement, this data becomes too complex to provide any solution.

To handle the barrier, we introduce a practical assumption: we simplify our knowledge about countries' characteristics and allocate them among a limited number of groups. Let us now assume players can be split into two groups.

We consider the world of N_1 nations of type 1 and N_2 nations of type 2. We can say that parameters with indexes 1 and 2 determine the belonging of ith participant to the group 1 or 2, respectively. Current abatement benefit is assumed to be identical for all participants and depend on current total abatement Q

$$B(Q) = \frac{b}{N_1 + N_2}(aQ - \frac{1}{2}Q^2).$$

Marginal benefit function is

$$MB(Q) = b(a - Q)/(N_1 + N_2),$$

and global marginal benefit

$$(N_1 + N_2)MB(Q) = b(a - Q).$$

For the country of type i, the abatement cost function is assumed to be given by 3.

Net benefit of ith country is

$$\pi_i\left(\mathbf{q}^S, \mathbf{q}^F\right) = B(Q) - C_i(q_i^{S(F)}) \tag{4}$$

and global net benefit is

$$\Pi\left(\mathbf{q}^S, \mathbf{q}^F\right) = \underbrace{\left(n_1\pi_1(\mathbf{q}^S, \mathbf{q}^F) + n_2\pi_2(\mathbf{q}^S, \mathbf{q}^F)\right)}_{\text{players from } S}$$

$$+ \underbrace{\left((N_1 - n_1)\pi_1(\mathbf{q}^S, \mathbf{q}^F) + (N_2 - n_2)\pi_2(\mathbf{q}^S, \mathbf{q}^F)\right)}_{\text{players from } F} \tag{5}$$

$$= (N_1 + N_2)B(Q) + n_1C_1(q_1^S) + n_2C_2(q_2^S) + (N_1 - n_1)C_1(q_1^F) + (N_2 - n_2)C_2(q_2^F).$$

3.2.1 Full Cooperation

First, we assume that the grand coalition $S = \mathcal{N}$ has been formed. Under full cooperation, members of the grand coalition maximize the joint net benefit $\Pi(\mathbf{q^c})$,

$$\Pi(\mathbf{q^c}) = b\left(aQ_c - \frac{1}{2}Q_c^2\right) - \left[\frac{1}{2}c_1N_1\,(q_1^c)^2 + \frac{1}{2}c_2N_2\,(q_2^c)^2\right],$$

where q_i^c is an abatement effort chosen by a signatory of type i, Q_c is aggregate abatement achieved by the grand coalition $\mathcal{N}$. The first order conditions

$$\begin{cases} \partial\Pi(\mathbf{q^c})/\partial q_1^c = 0, & \text{for } N_1 \text{ players of type 1,} \\ \partial\Pi(\mathbf{q^c})/\partial q_2^c = 0, & \text{for } N_2 \text{ players of type 2,} \end{cases}$$

deliver

$$\begin{cases} b(a - Q_c) - c_1 q_1^c = 0, & \text{for } N_1 \text{ players of type } 1, \\ b(a - Q_c) - c_2 q_2^c = 0, & \text{for } N_2 \text{ players of type } 2. \end{cases}$$

The first order condition requires setting each country's marginal cost of abatement $MC_i(q_i^c), i \in 1,2$, equal to the global marginal benefit of abatement $\sum MB(Q_c)$. Thus obtained system

$$\begin{cases} c_1 q_1^c = b(a - N_1 q_1^c - N_2 q_2^c), \\ c_2 q_2^c = b(a - N_1 q_1^c - N_2 q_2^c) \end{cases}$$

yields the aggregate emission level under full cooperation

$$Q_c = N_1 q_1^c + N_2 q_2^c = \frac{a(N_1 + pN_1)}{\gamma + N_1 + \zeta N_2}.$$

Individual abatements are

$$q_1^c = \frac{a}{\gamma + N_1 + \zeta N_2}, \text{ for each of } N_1 \text{ countries of type } 1,$$

$$q_2^c = p q_1^c = \frac{ap}{\gamma + N_1 + \zeta N_2}, \text{ for each of } N_2 \text{ countries of type } 2,$$

where $\gamma = c_1/b$ and $\zeta = c_1/c_2$.

3.2.2 Non-Cooperative Case

In the non-cooperative case all players act as singletons playing Nash strategies against each other. Each country of type i, $i = 1,2$, chooses the abatement level q_i^o so that to maximize its payoff, taking the other countries' abatements as given. That is, each of $\mathcal{N}$ players behaves in a typical Nash-Cournot fashion maximizing its net benefit $\pi_i(\mathbf{q^o}), i = 1,2$,

$$\pi_i(\mathbf{q^o}) = \frac{b}{N_1 + N_2}\left(aQ_o - \frac{1}{2}Q_o^2\right) - \frac{1}{2}c_i\,(q_i^o)^2,$$

where q_i^o is abatement effort of a singleton of type i and Q_o is aggregate emission reduction achieved in pure non-cooperative case. The first order conditions

$$\frac{\partial \pi_i(\mathbf{q^o})}{\partial q_i^o} = 0, \quad i = 1,2,$$

deliver

$$\begin{cases} \frac{b}{N_1+N_2}(a - Q_o) - c_1 q_1^o = 0, & \text{for } N_1 \text{ players of type } 1, \\ \frac{b}{N_1+N_2}(a - Q_o) - c_2 q_2^o = 0, & \text{for } N_2 \text{ players of type } 2 \end{cases}$$

and require setting each country's own marginal benefit $MB(Q_o)$ equal to its own marginal cost of abatement $MC_i(q_i^o)$. We obtain the following system

$$\begin{cases} c_1 q_1^o = \frac{b}{N_1+N_2}(a - N_1 q_1^o - N_2 q_2^o), \\ c_2 q_2^o = \frac{b}{N_1+N_2}(a - N_1 q_1^o - N_2 q_2^o). \end{cases}$$

The solution is

$$q_1^o = \frac{a}{\gamma(N_1+N_2)+(N_1+\zeta N_2)}, \text{ for each of } N_1 \text{ countries from group 1,}$$
$$q_2^o = \zeta q_1^o = \frac{a\zeta}{\gamma(N_1+N_2)+(N_1+\zeta N_2)}, \text{ for each of } N_2 \text{ countries from group 2.}$$

Global abatement Q_o under non-cooperative behavior is

$$Q_o = \frac{a(N_1+\zeta N_2)}{(N_1+\zeta N_2)+\gamma(N_1+N2)}.$$

3.2.3 Environmental and Economic Benefit

It is easily verifiable that each country abates more and is better off in the case of full cooperation than under non-cooperation, *i.e.* $q_1^c > q_1^o, \quad q_2^c > q_2^o$ and $\pi_i(\mathbf{q}^o) > \pi_i(\mathbf{q}^c)$.

The environmental benefit of cooperation is

$$Q_c - Q_o = \frac{a\gamma(N_1+\zeta N_2)(N_1+N_2-1)}{(\gamma+N_1+\zeta N_2)(N_1+N_1\gamma+\zeta N_2+\gamma N_2)}.$$

Let us introduce notation $\alpha = N_1 + N_2$ and $\beta = N_1 + \zeta N_2$. Then

$$Q_c - Q_o = \frac{a\gamma\beta(\alpha-1)}{(\gamma+\beta)(\gamma\alpha+\beta)}. \tag{6}$$

The economic benefit is

$$\Pi(\mathbf{q}^c) - \Pi(\mathbf{q}^o) = \frac{1}{2}\left[\frac{(N_1^3+2N_1^2N_2-2N_1^2+\zeta N_1^2N_2+N_1-2N_1N_2}{(N_1+\gamma N_1+\zeta N_2+\gamma N_2)^2(\gamma+N_1+\zeta N_2)}\right. \tag{7}$$
$$\left.+\frac{N_1N_2^2-2\zeta N_1N_2+2\zeta N_1N_2^2+\zeta N_2-2\zeta N_2^2+\zeta N_2^3)c_1\gamma a^2}{(N_1+\gamma N_1+\zeta N_2+\gamma N_2)^2(\gamma+N_1+\zeta N_2)}\right].$$

Lemma 3.2.1 *Let*

$$\theta(N_1,N_2) = \frac{-(N_1+N_2)+\sqrt{(N_1+N_2)^2+8(N_2+N_1)}}{4(N_1+N_2)(N_1+\zeta N_2)}$$
$$= \frac{-\alpha+\sqrt{\beta^2+8\alpha}}{4\alpha\beta},$$
$$\vartheta(N_1,N_2) = \frac{-N_1-\frac{1}{4}\gamma+\frac{1}{4}\sqrt{\gamma^2+8\gamma^2N_2+8\gamma^2N_1}}{N_2}.$$

Then

(*i*) $\Pi(\mathbf{q}^c) - \Pi(\mathbf{q}^o)$ *increases in b, when* $0 < \gamma < \theta(N_1,N_2)$ *and decreases if* $\gamma > \theta(N_1,N_2)$*;*

(*ii*) $\Pi(\mathbf{q}^c) - \Pi(\mathbf{q}^o)$ *increases in* c_1*;*

(iii) *$\theta(N_1, N_2)$ is monotonic and approaches values of 1 and ζ as N_1 and N_2 becomes very large respectively;*

($iiii$) *$\Pi(\mathbf{q}^c) - \Pi(\mathbf{q}^o)$ decreases in ζ, when*

$$\zeta > \max\{0, \vartheta(N_1, N_2)\}.$$

P r o o f.

(i) Net benefit difference between full cooperation and pure non-cooperative case is given by (7). Differentiating it with respect to γ (while c_1 is constant) delivers that

$$\frac{\partial\left(\Pi(\mathbf{q}^c) - \Pi(\mathbf{q}^o)\right)}{\partial\gamma} \gtreqqless 0$$
$$\Updownarrow$$
$$2(N_1+N_2)^2\gamma^3 + (N_1+\zeta N_2)(N_1+N_2)^2\gamma^2$$
$$-2(N_1+\zeta N_2)^2(N_1+N_2)\gamma - (N_1+\zeta N_2)^3 \lesseqqgtr 0. \qquad (8)$$

The cubic equation has three roots among which only $\gamma = \theta(N_1, N_2)$ is positive. Expression (8) is positive when $\gamma > \theta(N_1, N_2)$.

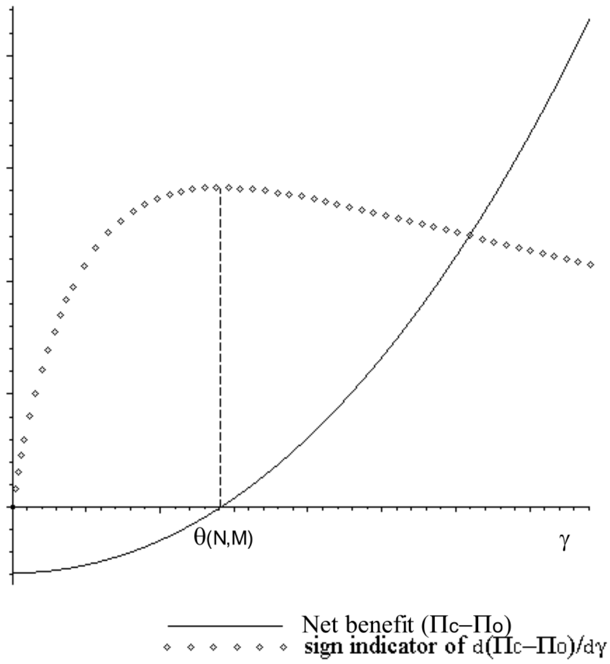


FIGURE 1 Dependence of the net benefit on parameter γ.

(ii) From (8) it follows immediately that $\partial(\Pi(\mathbf{q}^c) - \Pi(\mathbf{q}^o))/\partial c_1 \,|_{\gamma=const} > 0$.

(iii) Finding the limits,

$$\lim_{N_1\to\infty} \frac{-(N_1+N_2)+\sqrt{(N_1+N_2)^2+8(N_2+N_1)}}{4(N_1+N_2)(N_1+\zeta N_2)} = 1,$$

$$\lim_{N_2\to\infty} \frac{-(N_1+N_2)+\sqrt{(N_1+N_2)^2+8(N_2+N_1)}}{4(N_1+N_2)(N_1+\zeta N_2)} = \zeta.$$

Coefficient ζ determines c_1/c_2, so it is essential that it belongs to $(0, 1]$.

$(iiii)$ Solution of equation

$$\frac{\partial(\Pi(\mathbf{q}^c)-\Pi(\mathbf{q}^o))}{\partial\zeta} = 0$$

consists of three roots, among which only one can be positive,

$$\zeta = \vartheta(N_1, N_2).$$

When $\zeta > \max\{0, \vartheta(N_1, N_2)\}$, partial differential is negative.

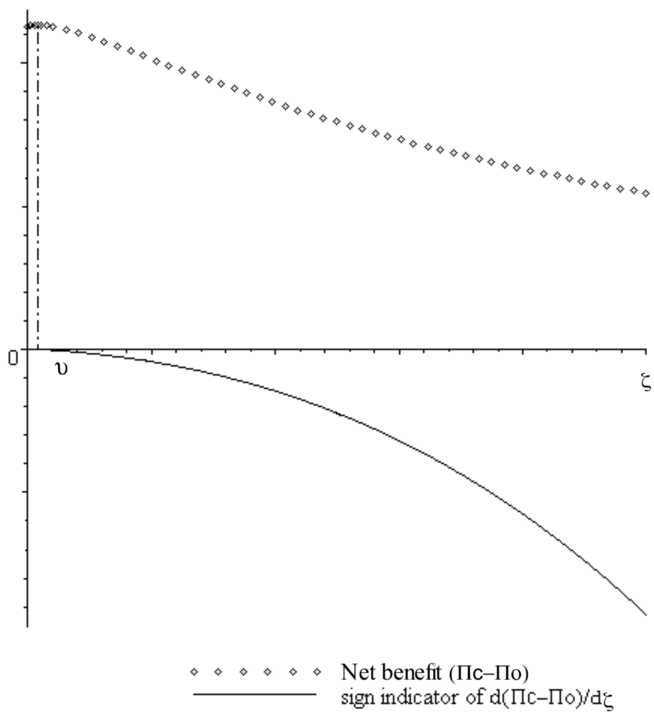


FIGURE 2 Dependence of the net benefit on parameter ζ.

Value ϑ can be positive only if inequality holds

$$2N_1^2 + N_1\gamma \le \gamma^2(N_1+N_2),$$

it would imply that number of countries N_1 from group 1 is relatively small in comparison to N_2. ■

In a similar way to [Barrett 1994a], we can make the following conclusion from Lemma 3.2.1:

- the gains of cooperation $\Pi(\mathbf{q}^c) - \Pi(\mathbf{q}^o)$ are larger the closer γ to $\theta(N_1, N_2)$, the closer ζ to $\max\{0, \vartheta\}$;
- when p is small, *i.e.* c_2 is much bigger than c_1 (marginal abatement cost of group 2 grows faster than those of 1) then the gains of cooperation $\Pi(\mathbf{q}^c) - \Pi(\mathbf{q}^o)$ are large;
- when c_1 is small and b is large, the gains of cooperation are relatively small, and countries will have slight incentives to join the agreement;
- when c_1 is large and b is small then difference $Q_c - Q_o$ is relatively small and ecological benefit is not sufficient;
- when $c_1 \approx b$ and they are small then $Q_c - Q_o$ is large, but $\Pi(\mathbf{q}^c) - \Pi(\mathbf{q}^o)$ is small;
- when $c_1 \approx b$ and they are large then both $Q_c - Q_o$ and $\Pi(\mathbf{q}^c) - \Pi(\mathbf{q}^o)$ are large.

The last two statements describe a set of parameters, where stable coalition might form.

3.2.4 Formation of Self-Enforcing Coalition

Let us consider formation of a coalition S, which is composed of n_1 and n_2 countries from groups 1 and 2, which make the decision to sign the IEA. There are then $(n_1 + n_2)$ signatories and $(N_1 - n_1) + (N_2 - n_2)$ free-riders. To determine a stable structure $\mathbf{n} = (n_1, n_2)$ of the coalition S we apply the principle of internal and external stability of a coalition, also known as *self-enforcing*, [d'Aspremont *et al.* 1983].

Definition 3.2.1 *A coalition S, characterized by vector $\mathbf{n} = (n_1, n_2)$ of signatories, is self-enforcing if*

$$\pi_i(\mathbf{q}^{S\setminus\{i\}}, \mathbf{q}^{F\cup\{i\}}) \le \pi_i(\mathbf{q}^S, \mathbf{q}^F), \quad i \in S, \tag{9}$$

$$\pi_i(\mathbf{q}^{S\cup\{i\}}, \mathbf{q}^{F\setminus\{i\}}) \le \pi_i(\mathbf{q}^S, \mathbf{q}^F), \quad i \in F. \tag{10}$$

Inequality (9) sets condition of internal stability, *i.e.* no member of S, described by vector $\mathbf{n}$ prefers to withdraw from the agreement (so that coalition would be characterized by vector $(n_1 - 1, n_2)$ or $(n_1, n_2 - 1)$). Condition (10) of external stability guarantees that no free-rider from set F prefers to join the coalition S, thus increasing number of signatories (so that coalition structure would be specified by vector $(n_1 + 1, n_2)$ or $(n_1, n_2 + 1)$). Stability conditions ensure that no player unilaterally deviates.

Signatories of IEA reduce q_1^S, q_2^S of their emissions, and total abatement undertaken by coalition is

$$Q_S = \sum_{i=1}^{2} n_i q_i^S. \tag{11}$$

In a similar manner, each free-rider from groups 1 and 2 emits q_1^F and q_2^F, yielding total abatement level

$$Q_F = \sum_{i=1}^{2}(N_i - n_i)q_i^F.$$

Supposing that a non-empty coalition S, characterized by $\mathbf{n} = (n_1, n_2)$ of signatories, has formed. Let us determine optimal abatement strategies of players as Stackelberg equilibrium in the two level game $\Gamma_0(S)$, where the coalition S acts as a leader and free-riders from set F accept position of the followers. Thus free-riders reduce their emission non-cooperatively taking the choice of signatories into account. Every free-rider of type i, $i = 1, 2$, maximizes its net benefit non-cooperatively

$$\max_{q_i^F} \pi_i(\mathbf{q}^S, \mathbf{q}^F), \qquad i \in F,$$

where

$$\pi_i(\mathbf{q}^S, \mathbf{q}^F) = \frac{b}{N_1 + N_2}\left(a(Q_S + Q_F) - \frac{1}{2}(Q_S + Q_F)^2\right) - \frac{1}{2}c_i\left(q_i^F\right)^2.$$

Reaction function of $N_i - n_i$ free-riders from set F of type $i, i = 1, 2$ can be found from the equation

$$MC_i(q_i^F) = MB(Q),$$

that is equivalent to

$$\frac{b}{N_1 + N_2}(a - Q_s - Q_F) - c_i q_i^F = 0.$$

Thus we obtain

$$\begin{cases} q_1^F = (a - Q_S)/(\gamma(N_1 + N_2) + (N_1 - n_1) + \zeta(N_2 - n_2)), \\ q_2^F = \zeta q_1^F. \end{cases}$$

The reaction function of non-signatories is

$$Q_F = g(a - Q_S), \tag{12}$$

where

$$g = \frac{(N_1 - n_1) + \zeta(N_2 - n_2)}{\gamma(N_1 + N_2) + (N_1 - n_1) + \zeta(N_2 - n_2)}.$$

Signatories choose their abatement level by maximizing their collective net benefit while taking into account behavior of non-signatories. Abatement Q_S is chosen by solving the following constrained maximization problem

$$\begin{array}{l} \max \sum_{i \in S} \pi_i(\mathbf{q}^S, \mathbf{q}^F), \\ \text{subject to (12)}, \end{array}$$

where $\pi_i(\mathbf{q}^S, \mathbf{q}^F)$ is the net benefit function of each signatory of type i. Solution of maximization problem is

$$q_2^S = \zeta q_1^S = \frac{a\zeta(n_1 + n_2)(1 - g)^2}{\gamma(N_1 + N_2) + (n_1 + \zeta n_1)(n_1 + n_2)(1 - g)^2}.$$

According to (11) coalition of n_1 1-counties and n_2 2-counties undertake the following abatement

$$Q_S = \frac{a(n_1 + n_2\zeta)(n_1 + n_2)(1-g)^2}{\gamma(N_1 + N_2) + (n_1 + n_2\zeta)(n_1 + n_2)(1-g)^2}. \tag{13}$$

Substituting (13) into reaction function (12) of non-signatories abatement, we obtain the following abatement for free-riding $(N_1 + N_2 - n_1 - n_2)$ countries

$$Q_F = \frac{\gamma a g(N_1 + N_2)}{\gamma(N_1 + N_2) + (n_1 + n_2\zeta)(n_1 + n_2)(1-g)^2},$$

individual abatements of free-riders will be

$$q_1^F = \frac{Q_F}{(N_1 - n_1) + \zeta(N_2 - n_2)},$$

$$q_2^F = \frac{\zeta Q_F}{(N_1 - n_1) + \zeta(N_2 - n_2)}.$$

Total abatement is $Q = Q_S + Q_F$.

The remaining problem is to determine n_1 and n_2. We invoke Definition 3.2.1, applying stability conditions (9) and (10). The net benefits (1) for each signatory of the coalition S and each free-rider from set F need to be calculated

$$\pi_i(\mathbf{q}^S, \mathbf{q}^F) = B(Q) - C_i(q_i^S), \quad i \in S,$$

$$\pi_i(\mathbf{q}^S, \mathbf{q}^F) = B(Q) - C_i(q_i^F), \quad i \in F,$$

when the coalition S is described with vector $\mathbf{n} = (n_1, n_2)$ and set F is characterized by vector $(N_1 - n_1, N_2 - n_2)$. We should point out that values q_i^S, q_i^F and Q_S, Q_F depend on coalitional structure, and thus on vector $\mathbf{n}$. Hence to identify the stable structure of S, we substitute expressions of players' abatement strategies into the net benefit functions, and examine two scenarios:

1. a member of type 1 may deviate from the coalition or join the coalition; then we need to solve the following system of inequalities

$$\begin{aligned}&B(Q_S(n_1-1, n_2) + Q_F(n_1-1, n_2)) - C_1(q_1^F(n_1-1, n_2)) \leq\\ &B(Q_S(n_1, n_2) + Q_F(n_1, n_2)) - C_1(q_1^S(n_1, n_2)),\end{aligned}$$

$$\begin{aligned}&B(Q_S(n_1, n_2) + Q_F(n_1, n_2)) - C_1(q_1^F(n_1, n_2)) \geq\\ &B(Q_S(n_1+1, n_2) + Q_F(n_1-1, n_2)) - C_1(q_1^S(n_1, n_2));\end{aligned}$$

2. a member of type 2 may deviate from coalition or joins to coalition. System of inequalities is

$$\begin{aligned}&B(Q_S(n_1, n_2-1) + Q_F(n_1, n_2-1)) - C_2(q_2^F(n_1, n_2-1)) \leq\\ &B(Q_S(n_1, n_2) + Q_F(n_1, n_2)) - C_2(q_2^S(n_1, n_2)),\end{aligned} \tag{14}$$

$$\begin{aligned}&B(Q_S(n_1, n_2) + Q_F(n_1, n_2)) - C_2(q_2^F(n_1, n_2)) \geq\\ &B(Q_S(n_1, n_2+1) + Q_F(n_1, n_2+1)) - C_2(q_2^S(n_1, n_2)).\end{aligned} \tag{15}$$

Due to complexity of inequalities (14) and (15) in general case, we shall first go over to numerical simulations and then continue our analysis.

Example 3.2.1

Let us assume the following parameters of the model (see Table 2). We are go-

TABLE 2 Example 3.2.1 Model Parameters

$b = 1$
$a = 100$

$Type \quad 1$	$Type \quad 2$
$c_1 = 0.5$	$c_2 = 0.8$
$N_1 = 5$	$N_2 = 10$

ing to identify stable coalition structures, the net benefits, and abatement targets of the players. The following technique reveals self-enforcing structure (further analysis is based on data presented in Fig. 13 and 14 in Appendix 1). First, we fix one of the values, *e.g.* n_1, and run through all n_2 checking conditions (14) and (15). For example, let us assume $n_1 = 1$. After set of comparisons

$$\pi_2^F(1,0) = 271.825 < \pi_2^S(1,1) = 275.350,$$

$$\pi_2^F(1,1) = 274.293 < \pi_2^S(1,2) = 276.779,$$

process stops at $\pi_2^F(1,2) = 280.286 > \pi_2^S(1,3) = 279.747$, delivering external stability at $n_2^{ext} = 2$. Then we check internal stability likewise:

$$\pi_2^S(1,10) = 308.749 < \pi_2^F(1,9) = 326.130,$$

$$\pi_2^S(1,9) = 305.485 < \pi_2^F(1,8) = 323.083,$$
$$\pi_2^S(1,8) = 301.771 < \pi_2^F(1,7) = 318.879,$$
$$\pi_2^S(1,7) = 297.616 < \pi_2^F(1,6) = 313.262,$$
$$\pi_2^S(1,6) = 293.095 < \pi_2^F(1,5) = 306.113,$$
$$\pi_2^S(1,5) = 288.384 < \pi_2^F(1,4) = 297.643,$$
$$\pi_2^S(1,4) = 283.788 < \pi_2^F(1,3) = 288.596,$$
$$\pi_2^S(1,3) = 279.747 < \pi_2^F(1,2) = 280.286,$$
$$\pi_2^S(1,2) = 276.779 > \pi_2^F(1,1) = 274.293.$$

Internal stability is reached at $n_2^{int} = 2$. Since $n_2^{int} = n_2^{ext}$, which guarantees internal and external stability simultaneously, we are able to come to the second step. By a similar way we should fix now $n_2^* = 2$ and running through all n_1 we check if there is stability $n_1^* = n_1^{int} = n_1^{ext}$ and if so compare the obtained stable solution to the assumption n_1 made before searching for n_2^*. If n_1 is the same,

TABLE 3 Example 3.2.1 Self-Enforcing Coalitions

Structure	S_1 $n_1 = 0 \quad n_2 = 3$	S_2 $n_1 = 1 \quad n_2 = 2$
Coalition members' net benefit	– $\pi_2 = 275.986$	$\pi_1 = 272.226$ $\pi_2 = 276.779$
Sigantories' commitments	– $q_2^S = 4.301$	$q_1^S = 6.969$ $q_2^S = 4.355$
Free-riders' abatement	$q_1^F = 5.161$ $q_2^F = 3.22$	$q_1^F = 5.11$ $q_2^F = 3.194$
Emission reduction	$Q = 61.290$	$Q = 61.673$

then it means that self-enforcing equilibrium has been achieved, in a particular case $(n_1^* = 1, n_2^* = 2)$.

Under the given set of parameters, we obtain two self-enforcing coalitional structures of IEA, which are S_1 with $\mathbf{n}^* = (0,3)$ and S_2 with $\mathbf{n}^* = (1,2)$. Fig. 3 and 4 demonstrate graphical interpretation of the conditions of internal and external stability of both coalitions:
1) internal stability of S_1

$$\pi_2^S(0,3) = 275.986 > \pi_2^F(0,2) = 275.986,$$

2) external stability of S_1

$$\pi_2^F(0,3) = 279.223 > \pi_2^S(0,4) = 278.034,$$

$$\pi_1^F(0,3) = 276.726 > \pi_1^S(1,3) = 272.373,$$

3) internal stability of S_2

$$\pi_1^S(1,2) = 272.226 > \pi_1^F(0,2) = 271.721,$$

$$\pi_2^S(1,2) = 276.779 > \pi_2^F(1,1) = 274.293,$$

4) external stability of S_2

$$\pi_1^F(1,2) = 277.838 > \pi_1^S(2,2) = 274.121,$$

$$\pi_2^F(1,2) = 280.286 > \pi_2^S(1,3) = 279.747.$$

It is also important to note that each self-enforcing coalition structure brings different environmental benefit (see Table 3), *i.e.* $Q_{1^*,2^*} = 61.673$ and $Q_{0^*,3^*} = 61.290$. Thus from an environmental point of view, coalition structure $(n_1^* = 1, n_2^* = 2)$ is more beneficial. It directly follows from considering net benefits of signatories and free-riders

$$\pi_2^F(1^*,2^*) = 280.286 > \pi_2^F(0^*,3^*) = 279.223,$$

FIGURE 3 Example 3.2.1. Graphical analysis of a self-enforcing IEA, when $\mathbf{n}^* = (1, 2)$.

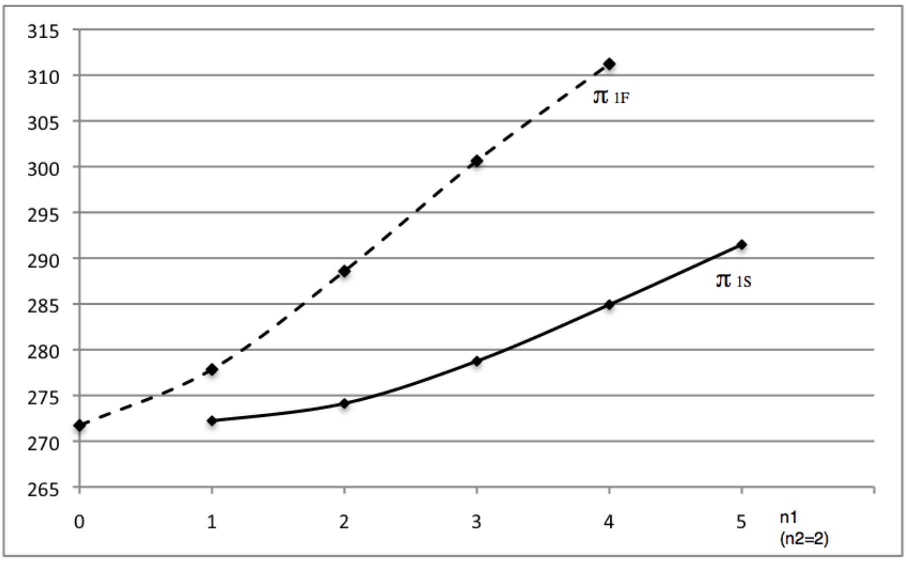
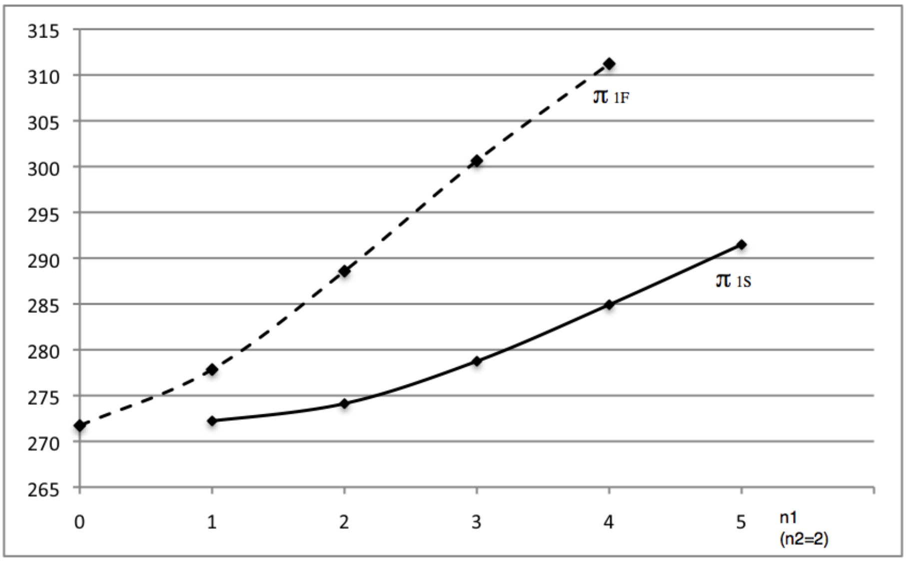


FIGURE 4 Example 3.2.1. Graphical analysis of self-enforcing IEA, when $\mathbf{n}^* = (0, 3)$.



$$\pi_2^S(1^*, 2^*) = 276.779 > \pi_2^S(0^*, 3^*) = 275.986,$$
$$\pi_1^F(1^*, 2^*) = 277.838 > \pi_1^F(0^*, 3^*) = 276.726,$$

Numerical examples, presented in Fig. 15 and 16 in Appendix 1, describe how high and low asymmetry of types and variation of initial parameters b, c_1 and c_2 effect the IEA structure. Despite the fact that full analytical characterization is not available due to complex model specification, these simulations reveal the relationship between parameters γ and ζ and the ability of IEA to fill the gap between full-cooperative and non-cooperative behaviors.

3.3 Coalition Formation Game among Heterogeneous Players of K Types

In this section we consider the coalition formation game among N heterogeneous players. Similar ideas of heterogeneous approaches can, for instance, be found in [Barrett 2001], where numerical simulations with $K = 2$ groups of players are performed, and where analytical analysis for $K = 2$ and $K = 3$ are presented. In addition to that, in [McGinty 2006] N pure asymmetric player game is presented. Further we present analysis of general case with K types of players in order to provide estimations of IEA size and structure and to study pollution transfers and rule of fair surplus allocation among IEA signatories.

Within the framework of this section we interpret agreement formation as a static two-level game $\Gamma_0(S) = \langle \mathcal{N}, \{q_i^S, q_i^F\}_{i=1}^K, \{\pi_i^S, \pi_i^F\}_{i=1}^K \rangle$, where q_i^S, q_i^F are players' strategies and π_i^S, π_i^F are net benefits of players of type i from S (and $F = \mathcal{N} \setminus S$). In the current game $\Gamma_0(S)$ the coalition S is the leader, and free-riders F are the followers. Strategies (abatement targets) $(\mathbf{q}^S, \mathbf{q}^F)$ are feasible if $Q \leq a$.

3.3.1 Pure Non-Cooperative Outcome

As before in the pure non-cooperative case all players act as singletons choosing Nash strategies to play against each other. Each country chooses its emission reduction level assuming others' rational reaction.

Individual abatement $q_i = q_i^o$, $i = 1, \dots, K$, is determined in such a way that maximize individual net benefit $\pi_i(\mathbf{q^o})$

$$\pi_i(\mathbf{q^o}) = \frac{b}{N}(aQ_o - \frac{1}{2}Q_o^2) - \frac{1}{2}c_i\,(q_i^o)^2 .$$

Consequently, we need to find the solution of the first order conditions

$$\begin{cases} \dfrac{\partial \pi_i(\mathbf{q^o})}{\partial q_i^o} = 0, \\ i = 1, \dots, K. \end{cases}$$

This leads each country to set its own marginal cost of abatement MC_i equal to marginal abatement benefit MB

$$\begin{cases} MC_i(q_i^o) = MB(Q_o), \\ i = 1, \dots, K, \end{cases}$$

that is equivalent to

$$c_i q_i^o = \frac{b}{N}(a - \sum_{l=1}^{K} N_l q_l^o), \quad i = 1, \dots, K. \tag{16}$$

To present the solution of non-cooperative outcome, we introduce the following notations.

- Let $\lambda = (\lambda_1, \dots, \lambda_t, \dots, \lambda_K)$ be a vector, where

$$\lambda_i = \frac{b}{c_i}. \tag{17}$$

- $\mathbf{N} = (N_1, \dots, N_i, \dots, N_K)$.
- Let $\bar{\mathbf{1}} = (1, \dots, 1)$ be a vector of units.
- For two given vectors $x = (x_1, \dots, x_r)$ and $y = (y_1, \dots, y_r)$ expression (x, y) means their scalar product and equals to $\sum_{l=1}^{r} x_i y_i$.

It follows from system (16) that

$$\begin{cases} q_1^o = \lambda_1 \frac{1}{N}(a - \sum_{l=1}^{K} N_l q_l^o), \\ \dots \\ q_i^o = \lambda_i \frac{1}{N}(a - \sum_{l=1}^{K} N_l q_l^o), \\ \dots \\ q_K^o = \lambda_K \frac{1}{N}(a - \sum_{l=1}^{K} N_l q_l^o). \end{cases}$$

It is easy to notice correlation $q_l^o = \lambda_l \frac{1}{\lambda_i} q_i^o$. Thus we obtain

$$\begin{cases} q_1^o = \lambda_1 \frac{1}{N}(a - \frac{q_1^o}{\lambda_1} \sum_{l=1}^{K} N_l \lambda_l), \\ \dots \\ q_i^o = \lambda_i \frac{1}{N}(a - \frac{q_i^o}{\lambda_i} \sum_{l=1}^{K} N_l \lambda_l), \\ \dots \\ q_K^o = \lambda_K \frac{1}{N}(a - \frac{q_K^o}{\lambda_K} \sum_{l=1}^{K} N_l \lambda_l). \end{cases}$$

It follows that

$$\begin{cases} q_1^o(N + \sum_{l=1}^{K} N_l \lambda_l) = \lambda_1 a, \\ \dots \\ q_i^o(N + \sum_{l=1}^{K} N_l \lambda_l) = \lambda_i a, \\ \dots \\ q_K^o(N + \sum_{l=1}^{K} N_l \lambda_l) = \lambda_K a. \end{cases}$$

Using notations described above, the non-cooperative solution can be presented as

$$q_i^o = \frac{a\lambda_i}{(\bar{1} + \lambda, \mathbf{N})}, \quad i = 1, \dots, K,$$

and total non-cooperative abatement

$$Q_o = \frac{a(\lambda, \mathbf{N})}{(\bar{1} + \lambda, \mathbf{N})}.$$

3.3.2 Full Cooperation

We need to assume that all players have made a decision of full cooperation $S = \mathcal{N}$ (*i.e.* the grand coalition is formed). To determine optimal abatements and corresponding net benefits of abatement, the grand coalition maximizes the global net benefit Π

$$\Pi(\mathbf{q^c}) = \sum_{i=1}^{K} N_i \pi_i(\mathbf{q^c}) = b(aQ_c - \frac{1}{2}Q_c^2) - \frac{1}{2}\sum_{i=1}^{K} N_i c_i\, (q_i^c)^2 .$$

Since players are homogeneous within one type, we can say that their strategies are also equal, thus q_i^c, $i = 1, \ldots, K$, denotes individual abatement of each country from group i under full cooperation. The problem

$$\max_{\mathbf{q^c}} \Pi(\mathbf{q^c})$$

leads to the fist order conditions

$$\frac{\partial \Pi(\mathbf{q^c})}{\partial q_i^c} = 0, \quad i = 1, \ldots, K.$$

We come to the following system

$$\begin{cases} N_1 b(a - \sum_{l=1}^{K} N_l q_l^c) - N_1 c_1 q_1^c = 0, \\ \ldots \\ N_i b(a - \sum_{l=1}^{K} N_l q_l^c) - N_i c_i q_i^c = 0, \\ \ldots \\ N_K b(a - \sum_{l=1}^{K} N_l q_l^c) - N_K c_K q_K^c = 0, \end{cases}$$

$$\begin{cases} c_1 q_1^c = b(a - \sum_{l=1}^{K} N_l q_l^c), \\ \ldots \\ c_i q_i^c = b(a - \sum_{l=1}^{K} N_l q_l^c), \\ \ldots \\ c_K q_K^c = b(a - \sum_{l=1}^{K} N_l q_l^c). \end{cases} \tag{18}$$

System (18) proves that in case of full cooperation each country's marginal cost of abatement is equal to the global marginal benefit of abatement

$$MC_i(q_i^c) = NMB(Q_c), \quad i = 1, \ldots, K.$$

Using correlation $q_j^c = \lambda_j \frac{1}{\lambda_i} q_i^c, j \neq i$, the system can be transformed to

$$\begin{cases} c_1 q_1^c = \lambda_1 (a - \frac{q_1}{\lambda_1} \sum_{l=1}^{K} N_l \lambda_l), \\ \ldots \\ c_i q_i^c = \lambda_i (a - \frac{q_i}{\lambda_i} \sum_{l=1}^{K} N_l \lambda_l), \\ \ldots \\ c_K q_K^c = \lambda_K (a - \frac{q_K}{\lambda_K} \sum_{l=1}^{K} N_l \lambda_l). \end{cases}$$

Commitments of each player of type i are determined as

$$q_i^c = \frac{a\lambda_i}{1 + (\lambda, \mathbf{N})}.$$

The global abatement is

$$Q_c = \frac{a(\lambda, \mathbf{N})}{1 + (\lambda, \mathbf{N})}.$$

3.3.3 Self-Enforcing Coalition Formation

In the framework, when participation decisions are made simultaneously, players have an incentive to cheat on the agreement and be better off by emission reduction achieved by other players who joined the IEA (see Example 3.2.1).

Let set S $(\varnothing \neq S \subseteq \mathcal{N})$ be a coalition of players that jointly intend to reduce their emissions, the rest of the players (free-riders) belong to the set $F = \mathcal{N} \setminus S$. Players simultaneously and voluntarily decide to join the coalition S or act independently. Denote n_i $(n_i \leq N_i, \quad i = 1, \ldots, K)$ as the number of players of type i that joined the agreement. They choose their abating strategies to maximize the net benefit of the coalition. The remaining $N_i - n_i$ players (free-riders) adjust their abatement levels non-cooperatively, maximizing individual net benefit. The vector $\mathbf{n} = (n_1, \ldots, n_i, \ldots, n_K)$ describes the IEA structure.

To determine players strategies $\left(\mathbf{q}^S, \mathbf{q}^F\right)$ in the two-level game $\Gamma_0(S)$ we apply the Stackelberg equilibrium concept. It implies that there is a Stackelberg leader (a coalition S of signatories) who takes into account the optimal choice of non-signatories (from the set F) which accept a position of the Stackelberg followers. In this case the agreement participants have an advantage towards free-riders as they choose their abatement levels based on the reaction functions of non-signatories. Free-riders are assumed to play non-cooperatively against each other and the coalition by choosing Nash strategies.

Lemma 3.3.1 *In the two level game $\Gamma_0(S)$ the Stackelberg equilibrium is unique and is constituted by the following strategies*

$$q_i^S = \frac{a\lambda_i(1-g)^2(\bar{1}, \mathbf{n})}{(\bar{1}, \mathbf{N}) + (1-g)^2(\bar{1}, \mathbf{n})(\lambda, \mathbf{n})}, \quad i = 1, \ldots, K, \tag{19}$$

$$q_i^F = \frac{\lambda_i a(\bar{1}, \mathbf{N})}{[(\bar{1}, \mathbf{N}) + (1-g)^2(\bar{1}, \mathbf{n})(\lambda, \mathbf{n})][(\bar{1} + \lambda, \mathbf{N}) - (\lambda, \mathbf{n})]}, \quad i = 1, \ldots, K, \tag{20}$$

where

$$g = \frac{(\lambda, \mathbf{N} - \mathbf{n})}{N + (\lambda, \mathbf{N} - \mathbf{n})}.$$

P r o o f. We are going to construct the Stackelberg equilibrium in the two level game $\Gamma_0(S)$. Assuming that the coalition S has chosen some feasible strategies q_i^S, free-riders of type i (from the set F) adjust their optimal abating efforts q_i^F by maximizing individual net benefit

$$\max_{q_i^F} \pi_i(\mathbf{q}^S, \mathbf{q}^F), \quad i = 1, \ldots, K,$$

which is expressed by

$$\pi_i(\mathbf{q}^S, \mathbf{q}^F) = \frac{b}{N}\left(a(Q_S + Q_F) - \frac{1}{2}(Q_S + Q_F)^2\right) - \frac{1}{2}c_i\left(q_i^F\right)^2.$$

The first order condition

$$\frac{\partial \pi_i(\mathbf{q}^S, \mathbf{q}^F)}{\partial q_i^F} = 0, \quad i = 1, \ldots, K,$$

leads to the system

$$\begin{cases} c_1 q_1^F = \frac{b}{N}\left(a - (Q_S + \sum_{l=1}^K q_l^F (N_l - n_l))\right), \\ \ldots \\ c_i q_i^F = \frac{b}{N}\left(a - (Q_S + \sum_{l=1}^K q_l^F (N_l - n_l))\right), \\ \ldots \\ c_K q_K^F = \frac{b}{N}\left(a - (Q_S + \sum_{l=1}^K q_l^F (N_l - n_l))\right). \end{cases} \tag{21}$$

Since $q_l^F = \lambda_l \frac{1}{\lambda_i} q_i^F$, the system (21) can be presented as

$$\begin{cases} q_1^F \left(N + \sum_{l=1}^K \lambda_l (N_l - n_l)\right) = \lambda_1 (a - Q_S), \\ \ldots \\ q_i^F \left(N + \sum_{l=1}^K \lambda_l (N_l - n_l)\right) = \lambda_i (a - Q_S), \\ \ldots \\ q_K^F \left(N + \sum_{l=1}^K \lambda_l (N_l - n_l)\right) = \lambda_K (a - Q_S). \end{cases}$$

The solution of the system is

$$\begin{aligned} q_i^F &= \frac{\lambda_i (a - Q_S)}{N + \sum_{l=1}^K \lambda_l N_l - \sum_{l=1}^K \lambda_l n_l} \\ &= \frac{\lambda_i (a - Q_S)}{(\bar{1} + \lambda, \mathbf{N}) - (\lambda, \mathbf{n})}, \end{aligned} \tag{22}$$

and since

$$\frac{\partial^2 \pi_i(\mathbf{q}^S, \mathbf{q}^F)}{\partial^2 q_i^F} = -\frac{b}{N} - c_i < 0,$$

the solution given in (22) is maximum and determines optimal strategies of the followers (free-riders' individual abatements). The total abatement of the free-riders is

$$Q_F = g(a - Q_S), \tag{23}$$

where

$$g = \frac{(\lambda, \mathbf{N} - \mathbf{n})}{N + (\lambda, \mathbf{N} - \mathbf{n})}.$$

The expressions (22) and (23) can be interpreted as a rational (Nash equilibrium) reply of the followers to any of the leader's strategies. Taking the reaction functions (22) and (23) into account, the leader chooses its abatement Q_S by maximizing its aggregate net benefit

$$\max_{q_i^S} \sum_{i=1}^{K} n_i \pi_i(\mathbf{q}^S, \mathbf{q}^F), \tag{24}$$

where

$$\sum_{i=1}^{K} n_i \pi_i(\mathbf{q}^S, \mathbf{q}^F) = \frac{\sum_{i=1}^{K} n_i}{N} b \left(a(Q_S + g(a - Q_S)) - \frac{1}{2}(Q_S + g(a - Q_S))^2 \right)$$
$$- \frac{1}{2} \sum_{i=1}^{K} n_i c_i \left(q_i^S \right)^2.$$

The maximization problem (24) leads to the first order conditions

$$\begin{cases} \partial \sum_{i=1}^{K} n_i \pi_i(\mathbf{q}^S, \mathbf{q}^F) / \partial q_i^S = 0, \\ i = 1, \ldots, K. \end{cases}$$

Thus we come to the system

$$\begin{cases} \frac{\sum_{l=1}^{K} n_l}{N} b \left(a(1-g)n_1 - (1-g)(Q_S(1-g) + ag)n_1 \right) - c_1 n_1 q_1^S = 0, \\ \ldots \\ \frac{\sum_{l=1}^{K} n_l}{N} b \left(a(1-g)n_i - (1-g)(Q_S(1-g) + ag)n_i \right) - c_i n_i q_i^S = 0, \\ \ldots \\ \frac{\sum_{l=1}^{K} n_l}{N} b \left(a(1-g)n_K - (1-g)(Q_S(1-g) + ag)n_K \right) - c_K n_K q_K^S = 0, \end{cases}$$

$$\begin{cases} c_1 q_1^S = \frac{\sum_{l=1}^{K} n_l}{N} b \left(a(1-g)^2 - Q_S(1-g)^2) \right), \\ \ldots \\ c_i q_i^S = \frac{\sum_{l=1}^{K} n_l}{N} b \left(a(1-g)^2 - Q_S(1-g)^2) \right), \\ \ldots \\ c_K q_K^S = \frac{\sum_{l=1}^{K} n_l}{N} b \left(a(1-g)^2 - Q_S(1-g)^2) \right). \end{cases}$$

Using the expression (17), system can be converted to

$$\begin{cases} N q_1^S = \lambda_1 \left(a(1-g)^2 - Q_S(1-g)^2 \right) \sum_{l=1}^{K} n_l, \\ \ldots \\ N q_i^S = \lambda_i \left(a(1-g)^2 - Q_S(1-g)^2 \right) \sum_{l=1}^{K} n_l, \\ \ldots \\ N q_K^S = \lambda_K \left(a(1-g)^2 - Q_S(1-g)^2 \right) \sum_{l=1}^{K} n_l. \end{cases}$$

Applying correlation among individual abatement levels $q_i^S = \lambda_i \frac{1}{\lambda_l} q_l^S$, the system can be presented as

$$\begin{cases} N q_1^S = \lambda_1 \left(a(1-g)^2 - (1-g)^2 \frac{1}{\lambda_1} q_1^S \sum_{l=1}^{K} n_l \lambda_i \right) \sum_{l=1}^{K} n_l, \\ \ldots \\ N q_i^S = \lambda_i \left(a(1-g)^2 - (1-g)^2 \frac{1}{\lambda_i} q_i^S \sum_{l=1}^{K} n_l \lambda_i \right) \sum_{l=1}^{K} n_l, \\ \ldots \\ N q_K^S = \lambda_K \left(a(1-g)^2 - (1-g)^2 \frac{1}{\lambda_K} q_K^S \sum_{l=1}^{K} n_l \lambda_i \right) \sum_{l=1}^{K} n_l, \end{cases}$$

which is equivalent to

$$\begin{cases} q_1^S \left(N + (1-g)^2 \sum_{l=1}^K (n_l \lambda_l) \sum_{l=1}^K n_l \right) = \lambda_1 a (1-g)^2 \sum_{l=1}^K n_l, \\ \dots \\ q_i^S \left(N + (1-g)^2 \sum_{l=1}^K (n_l \lambda_l) \sum_{l=1}^K n_l \right) = \lambda_i a (1-g)^2 \sum_{l=1}^K n_l, \\ \dots \\ q_K^S \left(N + (1-g)^2 \sum_{l=1}^K (n_l \lambda_l) \sum_{l=1}^K n_l \right) = \lambda_K a (1-g)^2 \sum_{l=1}^K n_l. \end{cases}$$

The solution of the maximization problem (24) is

$$q_i^S = \frac{a\lambda_i (1-g)^2 (\bar{1}, \mathbf{n})}{(\bar{1}, \mathbf{N}) + (1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})}, \quad i = 1, \dots, K,$$

since

$$\frac{\partial^2 \sum_{i=1}^K n_i \pi_i (\mathbf{q}^S, \mathbf{q}^F)}{\partial^2 q_i^S} = -\frac{b}{N} - c_i < 0.$$

The aggregate coalitional abatement is given by

$$Q_S = \frac{a(1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})}{(\bar{1}, \mathbf{N}) + (1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})}. \tag{25}$$

The solution $(\mathbf{q}^S, \mathbf{q}^F)$ is feasible because

$$\begin{aligned} Q &= Q_S + Q_F = Q_S + g(a - Q_S) \\ &= a \frac{(1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})}{(\bar{1}, \mathbf{N}) + (1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})} + a \frac{g(\bar{1}, \mathbf{N})}{(\bar{1}, \mathbf{N}) + (1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})} \\ &= a \left[\frac{(1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})}{(\bar{1}, \mathbf{N}) + (1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})} + \frac{g(\bar{1}, \mathbf{N})}{(\bar{1}, \mathbf{N}) + (1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})} \right] \\ &= a \left[1 - \frac{(1-g)(\bar{1}, \mathbf{N})}{(\bar{1}, \mathbf{N}) + (1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})} \right]. \end{aligned}$$

Here we have

$$1 - g = \frac{1}{1 + (\lambda, \mathbf{N} - \mathbf{n})} \in (0, 1]$$

and

$$\frac{(\bar{1}, \mathbf{N})}{(\bar{1}, \mathbf{N}) + (1-g)^2 (\bar{1}, \mathbf{n})(\lambda, \mathbf{n})} \in (0, 1].$$

As a result, $Q = Q_S + Q_F \leq a$. This completes the proof. ■

It is also important to point out that values $(\mathbf{q}^S, \mathbf{q}^F)$ are positive and finite because the parameters a, b and c_i, $i = 1, \dots, K$, are positive.

As before we assume that the stability of the coalition S is associated with a principle of a self-enforcement under the formal players' payoffs.

Definition 3.3.1 *A coalition S, characterized by vector $\mathbf{n}$ of signatories of $K \leq N$ types, is self-enforcing in the game $\Gamma_0(S)$, if for each type i, $i = 1, \dots, K$,*

$$\pi_i(\mathbf{q}^S, \mathbf{q}^F) \geq \pi_i(\mathbf{q}^{S \setminus \{i\}}, \mathbf{q}^{F \cup \{i\}}), \quad i \in S, \tag{26}$$

where $(\mathbf{q}^S, \mathbf{q}^F)$ *is the Stackelberg equilibrium in the game* $\Gamma_0(S)$ *and* $(\mathbf{q}^{S\setminus\{i\}}, \mathbf{q}^{F\cup\{i\}})$ *is the Stackelberg equilibrium in the game* $\Gamma_0(S \setminus \{i\})$,

$$\pi_i(\mathbf{q}^{S\cup\{i\}}, \mathbf{q}^{F\setminus\{i\}}) \leq \pi_i(\mathbf{q}^S, \mathbf{q}^F), \quad i \in F, \tag{27}$$

where $(\mathbf{q}^S, \mathbf{q}^F)$ *is the Stackelberg equilibrium in the game* $\Gamma_0(S)$ *and* $(\mathbf{q}^{S\cup\{i\}}, \mathbf{q}^{F\setminus\{i\}})$ *is the Stackelberg equilibrium in the game* $\Gamma_0(S \cup \{i\})$.

The inequality (26) guarantees internal stability of the coalition, *i.e.*, no member has reason to leave the IEA. The external stability condition (27) guarantees that no non-member prefers to join the coalition. In general, stability conditions ensure that no player benefits from unilateral deviation. In order to identify the structure of the self-enforcing coalition S, we substitute the abatement strategies $(\mathbf{q}^S, \mathbf{q}^F)$, presented in (19) and (20), into the conditions of internal/external stability, (26) and (27).

Since the benefit and cost functions (see Section 3.1) are nonlinear, and we deal with heterogeneous players, the system (26), (27) does not have an analytical solution. Numerical simulations can be found in [McGinty 2006] and [Dementieva & Pavlova 2007b]. It is shown that a solution of the system exists for a sufficiently large set of model parameters, and that the solution is often not unique, which means that one of a few coalitions can form.

To illustrate the concept of self-enforcing equilibrium we consider the following examples.

Example 3.3.1

Let us first consider parameters of the model as they were given in Example 3.2.1 (Table 2), where only two types of players are distinguished.

If players are split into two groups, it may be interpreted, for instance, as belonging to Annex B countries and non-Annex countries of the Kyoto Protocol. Under the given set of parameters, we have obtained two self-enforcing coalitional structures of IEA, *i.e.* $\mathbf{n} = \{(0,3), (1,2)\}$ (see Table 3). Table 3 demonstrates that each self-enforcing coalition structure brings different individual benefits to coalition members and different environmental benefit ($Q_{1^*,2^*} = 61.673$ and $Q_{0^*,3^*} = 61.290$). Thus coalition structure $\mathbf{n} = (1,2)$ is more preferable.

Example 3.3.2

This example demonstrates, how slight variation of model parameters may effect coalition formation. Let us consider a case similar to Example 3.3.1. Here we have strengthened the difference between players' marginal abatement costs. Emission reduction of the first type is characterized by moderate cost and abatements of the the second one are rather expensive. Other model parameters were left unchanged.

TABLE 4 Example 3.3.2 Model Parameters

$b = 1$
$a = 100$

Type 1	Type 2
$c_1 = 0.3$	$c_2 = 1$
$N_1 = 5$	$N_2 = 10$

We repeat the algorithm given in the previous example and obtain the following results. As Table 5 shows even slight variations of type characteristics reduces the number of stable coalitions. It appears that when the difference between nation types is increasing, countries with lower abatement cost are not motivated to join coalitions formed by countries with higher abatement costs, simply because it is less beneficial than free-riding.

TABLE 5 Example 3.3.2 Self-Enforcing Coalitions

	1
Structure	$n_1 = 0 \quad n_2 = 3$
Coalition members' net benefit	– $\pi_2 = 287.324$
Emission reduction	$Q = 64.42$

Example 3.3.3

Let us turn to a more general case and consider three types of players. As it mentioned before, it means that we distinguish between three groups: industrialized countries (like USA, European Union, Japan); rapidly developing counties (like Russia, China); and agricultural countries with low abatement cost.

TABLE 6 Example 3.3.3 Model Parameters

$b = 1$
$a = 100$

Type 1	Type 2	Type 3
$c_1 = 0.5$	$c_2 = 0.8$	$c_3 = 2$
$N_1 = 5$	$N_2 = 10$	$N_3 = 3$

This case is similar to Example 3.3.1. The simulation of stable coalition structure delivers:

Together with the previous example, results of these settings illustrate that nations with low abatement costs are not interested in joining agreements with other nations whose abatement costs are higher. Since the deal would not benefit them, they prefer to stay outside of the formed coalition. This and other undertaken numerical tests, [Barrett 1997a], [McGinty 2006], show that full cooperation is rather unlikely, coalition size is low and coalition structure lacks for diversity. Consequently, emission reduction is insufficient.

TABLE 7 Example 3.3.3 Self-Enforcing Coalitions

	1
Structure	$n_1 = 0 \quad n_2 = 0 \quad n_3 = 3$
Coalition members' net benefit	– $\pi_3 = 225.490$
Emission reduction	$Q = 57.647$

3.4 Environmental Efficiency

Definition 3.4.1 *An IEA, characterized by a coalition S, is environmentally efficient if for each $i \in S$*

$$Q \geq Q^{-i}. \tag{28}$$

Here Q is the total abatement undertaken by the players from the coalition S and free-riders from set F, as determined by (23) and (25) in the game $\Gamma_0(S)$. We denote by Q^{-i} the total abatement undertaken by the players from the coalition $S \setminus \{i\}$ and free-riders $F \cup \{i\}$, as determined in the game $\Gamma_0(S \setminus \{i\})$. The inequality (28) means that if any signatory of type i withdraws from S, it reduces the total abatement.

In the current section we assume that the players are symmetric. This assumption does not cause a major loss of generality and is an important step towards the more complex heterogeneous case. In this case, we deal with N homogeneous players with identical net benefit functions

$$\pi(\mathbf{q}^S, \mathbf{q}^F) = b(aQ - \frac{1}{2}Q^2) - \frac{1}{2}cq^2,$$

and $\lambda = b/c$, as in Section 3.1.

For convenience, we set

$$\eta_1, \eta_2 = \frac{(\lambda + N)(\lambda + 2N + 2\lambda N) \mp \sqrt{\lambda^2(\lambda + N)^2 + 4N^3(\lambda + N)(\lambda + 1)^2}}{2\lambda(\lambda + N)}. \tag{29}$$

Lemma 3.4.1 *The numbers η_1 and η_2 satisfy the following inequalities*

$$1 \leq \eta_1 \leq N,$$

and

$$\eta_2 > N.$$

P r o o f. Let us first prove that $\eta_1 < N$. This is equivalent to the inequality

$$\frac{\lambda^2 + 2N^2 + 3\lambda N - \sqrt{\lambda^2(\lambda + N)^2 + 4N^3(\lambda + N)(\lambda + 1)^2}}{2\lambda(\lambda + N)} < 0.$$

Since the denominator $2\lambda(\lambda+N)$ is positive, it suffices to show that

$$\lambda^2 + 2N^2 + 3\lambda N < \sqrt{\lambda^2(\lambda+N)^2 + 4N^3(\lambda+N)(\lambda+1)^2}.$$

As both sides of the inequality are positive, it is equivalent to show that

$$(\lambda^2 + 2N^2 + 3\lambda N)^2 < \lambda^2(\lambda+N)^2 + 4N^3(\lambda+N)(\lambda+1)^2.$$

From this, we obtain

$$-8\lambda N^3(N-1) - 4\lambda^3 N(N^2-1) - 8\lambda^2 N^2(N-1) - 4\lambda^2 N^2(N^2-1) < 0.$$

It is clear that the above inequality holds, and it follows that $\eta_1 < N$.

The other assertions follow similarly. ■

We denote by $\lceil x \rceil$ the smallest integer larger than $x \in \mathbb{R}$.

Theorem 3.4.1 *A coalition S, composed of n homogeneous players, is environmentally efficient if $n \geq \lceil \eta_1 \rceil$.*

P r o o f. In order to proof the statement, we introduce the following auxiliary function

$$f = Q - Q^{-i}, \quad i \in S.$$

Due to the symmetry of the players and formulas (19) and (20), it is easy to see that f is a function of $n = |S|$.

Equation $f(n) = 0$ has two solutions η_1 and η_2 (29). Lemma 3.4.1 provides us with properties of these solutions.

The agreement is feasible if $n \in [0, N]$. Since $n \leq 1$ corresponds to a trivial agreement structure, we assume that $n \in [2, N]$. By Lemma 3.4.1, η_1 is in the feasible interval $(1, N)$. The function $f(n)$ does not change sign on $(\eta_1, N]$, and so to prove that $f(n)$ is positive on this interval, it is sufficient to observe that

$$\begin{aligned} f(N) &= \frac{(\lambda N^2 + 2N^2 - 2N - \lambda)\lambda a}{(\lambda^2 + 3\lambda N + N^2 + N^3\lambda - 2\lambda N^2)(1+\lambda N)} \\ &= \frac{(\lambda(N^2-1) + 2N(N-1))\lambda a}{((\lambda+N)^2 + \lambda N(N-1)^2)(1+\lambda N)} > 0. \end{aligned}$$

Fig. 5 shows how the function $f(n)$ depends on λ when $N = 10$ and $a = 100$. It is easy to see that as λ grows, η_1 increases as well.

The coalition S is environmentally efficient if and only if $f(n) > 0$. Thus, the above argument shows that S is environmentally efficient when $n \geq \lceil \eta_1 \rceil$. ■

Theorem 3.4.1 means that an agreement needs to be formed by at least $\lceil \eta_1 \rceil$ participants to be environmentally efficient. This statement matches real-world examples where the ecological consistency of an agreement requires the participation of a certain number of actors (nations, countries, regions, *etc.*).

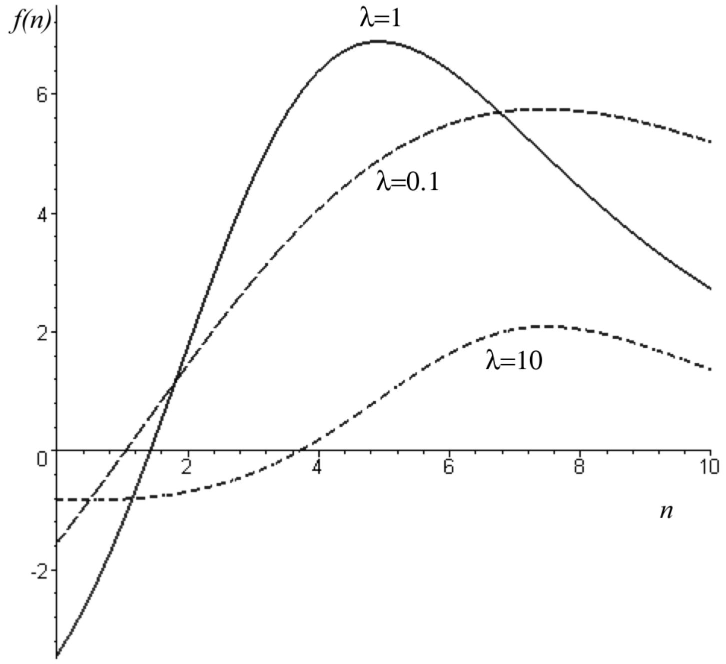


FIGURE 5 Function $f(n)$ under different settings. $N = 10, a = 100$.

Let us denote solution of the differential equation[1]

$$\frac{df(x)}{dx} = 0, \qquad x \geq 0,$$

as η^* (see Fig. 6).

If η^* belongs to the feasibility interval ($\eta^* \in [2, N]$), then we say that on the closed interval $[2, N]$ function $f(x)$ reaches its maximus at $x = \eta^*$. Otherwise if $\eta^* \notin [2, N]$, we say that function $f(x)$ gets its maximum value at $x = N$, and value η^* is reassigned as $\eta^* = N$.

We denote the closest integer to η^* by $[\eta^*]$. When the number of players n in the coalition is equal to $[\eta^*]$, the difference between Q and Q^{-i} is the largest.

Lemma 3.4.2 *A coalition S composed of $n = [\eta^*]$ homogeneous players is the most sensitive to withdrawal of its signatory in terms of environmental efficiency.*

Thus, the grand coalition of all N homogeneous players results in the largest total emission reduction.

---

[1] To provide consistency of the analysis, integer variable n is substituted by real variable x.

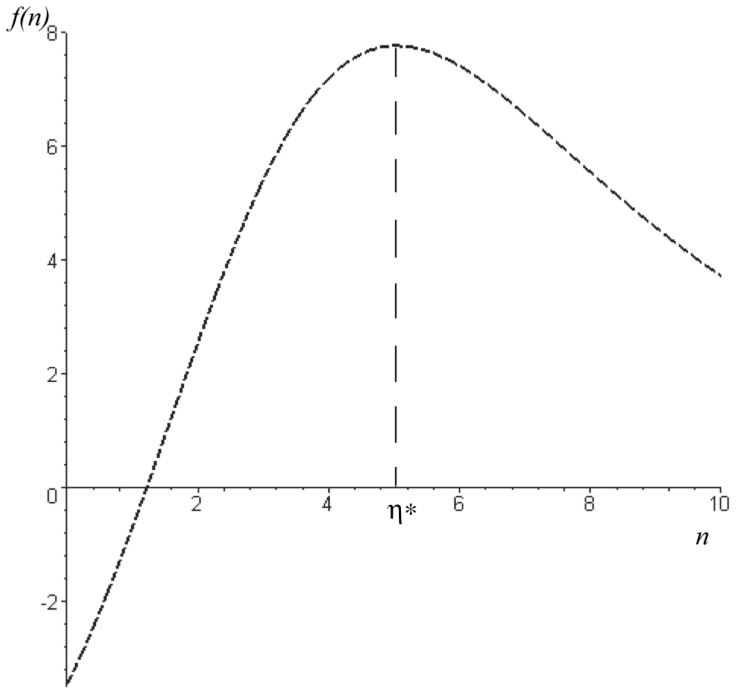


FIGURE 6 Maximum of Function $f(n)$. $N = 10, a = 100, \lambda = 0.5$.

Example 3.4.1

Let us reconstruct an example, given in [Barrett 1994a], which describes process of coalition formation among homogeneous players and results in a single self-enforcing IEA. According to [Barrett 1994a] parameters of the game $\Gamma_0(S)$ are as follows: $N = 10$, $a = 100$, $b = 1$ and $c = 0.25$. The process of identifying a self-enforcing coalition delivers a single stable IEA composed of 4 signatories.

This agreement, if formed, would provide total abatement $Q = 81.069$ (of $a = 100$), payoff of signatories and free-riders are $\pi_S = 472.16$ and $\pi_f = 474.913$, respectively.

Let us calculate $\eta_{1,2}$, determined in (29),

$$\eta_1 = 2.424, \qquad \eta_2 = 23.576.$$

According to Theorem 3.4.1 we can guarantee that the agreement is environmentally efficient when it is composed of $\lceil \eta_1 \rceil = 3$ and more players. From Table 8 one may see that for $n = 0, 1, 2$ $Q(n) - Q(n-1) < 0$ and when $n = 3, \ldots, N$ $Q(n) - Q(n-1) > 0$. Consequently, a coalition which consists of 4, satisfies this condition.

Lemma 3.4.2 says that a coalition is most sensitive, in terms of environmental efficiency, to withdrawal of signatory if it is composed of 6 players (here

TABLE 8 Example 3.4.1 Coaliton Stability Analysis.

n	q_i^S	q_i^F	π_S	π_F	Q	Π
0	0	8	0	472	80	4720
1	1.855	8.534	476.809	468.135	78.664	4690.022
2	4.158	8.732	474.012	466.643	78.170	4681.169
3	6.652	8.426	472.284	468.941	78.936	4699.436
4^*	8.909	7.572	472.160	474.913	81.069	4738.121
5	10.526	6.316	473.684	482.548	84.211	4781.163
6^{**}	11.342	4.915	476.371	489.431	87.713	4815.949
7	11.457	3.6	479.542	494.328	90.998	4839.776
8	11.096	2.497	482.663	497.273	93.759	4855.85
9	10.477	1.63	485.448	498.838	95.925	4867.872
10^{***}	9.756	0	487.805	0	97.56	4878.049

$\eta^* = 5.668$). It means that if the agreement of 5 (of 10) nations will be accessed by the 6th one, it will significantly improve abatement effort, from $Q = 84.211$ up to $Q = 87.713$. The biggest abatement achievable by the grand coalition is $Q = 97.56$.

Both theoretical results and numerical tests, presented in this chapter, show that full cooperation is rather unlikely and coalition structure lacks for diversity. Consequently, emission reduction is insufficient.

Under present circumstances we need to propose a mechanism, simple and comprehensive enough, which would allow one to reshape the agreement structure. One of the options is to introduce a rule, according to which coalition surplus is redistributed among possible coalition members. Such reallocation increases net benefits of potential signatories, who so far preferred to be outsiders, so that it would be more beneficial to access the agreement than to free-ride. On the other hand, according to this rule, welfare of current signatories will drop down but still will be high enough to keep coalition membership profitable. This rule can be represented, for instance, as side payments.

4 TRANSFER MECHANISMS

Since negotiators can anticipate how their disposition (to signatories and free-riders) is restructured by the agreement, they can design the agreement in such a way that it will contain an incentive mechanism system at the outcome. In Section 2.1 we have already described that incentive mechanisms incorporate methods of "stick and carrot", which mean transfers (financial transfers, emission trading) and punishments (trade restrictions *etc.*). In the present chapter we address the concept of potential self-enforcement (it requires weaker condition of internal stability), and suggest a transfer scheme interpreting it as emission trading.

4.1 Coalition Formation under Side Payments

Side payments aim to reallocate surplus, obtained by coalition S. Signatories of the IEA can share their individual surplus with other coalition members fully or partially (so that to keep at least a certain part of their individual surplus). We consider the first option. The total coalition surplus is

$$\Delta^S = \sum_{i=1}^{K} \Delta_i^S n_i, \tag{30}$$

where

$$\Delta_i^S = \pi_i(\mathbf{q}^S, \mathbf{q}^F) - \pi_i(\mathbf{q}^{S\setminus\{i\}}, \mathbf{q}^{F\cup\{i\}}), \quad i \in S, \tag{31}$$

where $\pi_i(\mathbf{q}^S, \mathbf{q}^F)$ is net benefit of signatory of type i, given structure $(n_1, \ldots, n_K)$ of coalition S; $\pi_i(\mathbf{q}^{S\setminus\{i\}}, \mathbf{q}^{F\cup\{i\}})$ is net benefit of a player of type i if it decides to free-ride from the agreement, so that coalition structure becomes $(n_1, \ldots, n_i - 1, \ldots, n_K)$.

If condition of internal stability (26) holds for n_i coalition members of type i, the correspondent term Δ_i^S is non-negative, otherwise it would be negative.

Surplus share can be considered in two cases.

1. The first case is when surplus share is introduced to reallocate profit among signatories of a self-enforcing coalition.

2. Surplus share can be introduced to attract more players to sign IEA; it is reasonable when size of self-enforcing coalition is rather small (and coalition abatement efforts are not sufficient). In this context, total surplus Δ^S consists of positive surplus provided by members of a self-enforcing IEA and non-positive one provided by outsiders. The coalition will grow in size, inviting new members of different types, until total surplus Δ^S becomes non-positive.

New individual payoff of signatory of the coalition S is

$$\sigma_i^S = \pi_i(\mathbf{q}^{S\setminus\{i\}}, \mathbf{q}^{F\cup\{i\}}) + \delta_i^S, \tag{32}$$

where $\delta_i^S \geq 0$ is such a surplus share of a signatory of type i that

$$\sum_{i=1}^{K} n_i \delta_i^S = \Delta^S. \tag{33}$$

Formula (32) means that each player, which has joined IEA, receives as much as it could get deviating from IEA, plus individual share of common surplus. Condition (33) says that the whole surplus is fully allocated.

Proposition 4.1.1 *Functions σ_i^S accurately reallocate the coalition gain.*

P r o o f. The proof of this statement is quite straightforward. We need to show that the sum of benefits received by coalition members under introduced side payments

$$\sum_{i=1}^{K} \sigma_i^S \cdot n_i$$

is equal to the total coalition gain received by signatories before introduction of side payments

$$\sum_{i=1}^{K} \pi_i(\mathbf{q}^S, \mathbf{q}^F) \cdot n_i.$$

To prove it, it is useful to represent the total gain as follows

$$\begin{aligned}\sum_{i=1}^{K} \sigma_i^S \cdot n_i &= \sum_{i=1}^{K} n_i \cdot \pi_i(\mathbf{q}^{S\setminus\{i\}}, \mathbf{q}^{F\cup\{i\}}) + \sum_{i=1}^{K} n_i \cdot \delta_i^S \\ &= \sum_{i=1}^{K} n_i \cdot \pi_i(\mathbf{q}^{S\setminus\{i\}}, \mathbf{q}^{F\cup\{i\}}) + \Delta^S.\end{aligned}$$

Using formulae (30) and (31), it follows that

$$\sum_{i=1}^{K} \sigma_i^S \cdot n_i = \sum_{i=1}^{K} \pi_i(\mathbf{q}^S, \mathbf{q}^F) \cdot n_i.$$

■

Hence condition of self-enforcing coalition can be rewritten as follows.

Definition 4.1.1 *A coalition S is called potentially internally stable if* $\Delta^S \geq 0$.

Definition 4.1.2 *A coalition S is called potentially self-enforcing if*

$$\Delta^S \geq 0,$$

$$\Delta^{S \cup \{i\}} \leq 0 \quad \text{for all } i = 1, \ldots, K.$$

Now we can go over to examples and consider game outcome when the given mechanism of side payments is allowed. Further we suppose that surplus Δ^S is distributed among coalition members according to the following allocation rule with weight coefficients

$$\delta_i^S = \frac{\lambda_i}{(\mathbf{n}, \lambda)} \Delta^S, \tag{34}$$

guaranteeing that individual share is characterized by individual marginal abatement. This allocation rule implies that nations with lower abatement costs receive greater surplus share than nations with higher abatement costs, and thus they can use surplus for investing in technological growth.

Example 4.1.1

We examine the model, where only two types of players are distinguished (see Examples 3.2.1 and 3.3.1).

TABLE 9 Example 4.1.1 Model Parameters

	Type 1	Type 2
$b = 1$	$c_1 = 0.5$	$c_2 = 0.8$
$a = 100$	$N_1 = 5$	$N_2 = 10$

Side payments will reshape coalition structure of self-enforcing agreement as it is illustrated in Table 10. It is easy to notice, that in comparison to results in Examples 3.2.1 and 3.3.1, self-enforcing coalitions doubled their number and structures became more diverse. It is also important to mark that payoffs of the second coalition have changed. Nation of type 1 receives $\pi_1 = 274.156$ (more than $\pi_1 = 272.226$ in case without side-payments). On the other hand, nations of type 2 lose by receiving $\pi_2 = 275.814$, which is less than $\pi_2 = 276.779$ in case when no side payments allowed. Despite that, two nations of type 2 are still willing to access agreement since it is more beneficial than leaving. The total surpluses that have been shared are

$$\Delta(n_1 = 1, n_2 = 2) = 6.33,$$

$$\Delta(n_1 = 2, n_2 = 1) = 5.478.$$

TABLE 10 Example 4.1.1 Self-Enforcing Coalitions

	1	2
Structure	$n_1 = 0 \quad n_2 = 3$	$n_1 = 1 \quad n_2 = 2$
Coalition members' payoff	– $\pi_2 = 275.986$	$\pi_1 = 274.156$ $\pi_2 = 275.814$
Emission reduction	$Q = 61.290$	$Q = 61.673$

	3	4
Structure	$n_1 = 2 \quad n_2 = 1$	$n_1 = 3 \quad n_2 = 0$
Coalition members' payoff	$\pi_1 = 273.979$ $\pi_2 = 275.753$	$\pi_1 = 273.916$ –
Emission reduction	$Q = 62.098$	$Q = 62.567$

Coalition $n_1 = 3$ and $n_2 = 0$ has become self-enforcing only after introduction of side payments. Before this, at least one nation of type 2 had incentive to join agreement of three countries of type 1, since it would increase its welfare. Let us assume that a player of type 2 attempts to access the coalition. This makes the agreement unstable because the joining of this player reduces individual payoffs of three former members of coalition.

Now when side payments have been started, player of type 2 has to share his surplus with three other members who suffer from his accession. It occurs that surplus is not big enough to cover losses of three nations of type 1 and that total coalition surplus is negative. From this moment player of type 2 places himself in an unprofitable situation (he loses from accessing the agreement) and has no incentive to make a claim for partnership.

Example 4.1.2

The second example demonstrates a situation with diverse players' parameters, when only single self-enforcing coalition is possible it may effect coalition formation. Let us consider a case similar to the first one.

TABLE 11 Example 4.1.2 Model Parameters

$b = 1$
$a = 100$

Type 1	*Type* 2
$c_1 = 0.3$	$c_2 = 1$
$N_1 = 5$	$N_2 = 10$

Side payments will reshape coalition structure of self-enforcing agreement as follows in Table 12.

TABLE 12 Example 4.1.2 Self-Enforcing Coalitions

	1	2
Structure	$n_1 = 0 \quad n_2 = 3$	$n_1 = 1 \quad n_2 = 2$
Coalition members' payoff	– $\pi_2 = 275.986$	$\pi_1 = 274.156$ $\pi_2 = 275.814$
Emission reduction	$Q = 61.290$	$Q = 61.673$

	3	4
Structure	$n_1 = 2 \quad n_2 = 1$	$n_1 = 3 \quad n_2 = 0$
Coalition members' payoff	$\pi_1 = 273.979$ $\pi_2 = 275.753$	$\pi_1 = 273.916$ –
Emission reduction	$Q = 62.098$	$Q = 62.567$

Example 4.1.3

Let us turn to the case when three types of players are involved and reconsider an already familiar example. Correspondent stable coalition structure can be found

TABLE 13 Example 4.1.3 Model Parameters

$b = 1$
$a = 100$

Type 1	*Type* 2	*Type* 3
$c_1 = 0.5$	$c_2 = 0.8$	$c_3 = 2$
$N_1 = 5$	$N_2 = 10$	$N_3 = 3$

in Table 14.

The presented examples demonstrate how introduction of side payments increases size of the coalition and structure diversity, even allowing abatement to grow to a certain degree. Side payments solve the problem of agreement inefficiency. At the same time another question arises. How to convince nations to share their welfare with others in order to increase agreement membership and accordingly reduce the level of global pollution? What is the suitable way of assigning commitments to potential signatories, so that side payments are naturally included into agreement patterns?

4.2 Emission Trading

In this section we propose one of the ways how emission trading can be designed. The bottleneck of this approach is to suggest sufficient a mechanism of setting the price and amount of tradable pollution permits (TPPs), as well as initial allocation of signatories' abatement commitments.

TABLE 14 Example 4.1.3 Self-Enforcing Coalitions

	1	2	3	4	5
Structure	$n_1 = 0$	$n_1 = 1$	$n_1 = 0$	$n_1 = 2$	$n_1 = 1$
	$n_2 = 0$	$n_2 = 0$	$n_2 = 1$	$n_2 = 0$	$n_2 = 1$
	$n_3 = 3$	$n_3 = 2$	$n_3 = 2$	$n_3 = 1$	$n_3 = 1$
Coalition	–	$\pi_1 = 222.479$	–	$\pi_1 = 222.193$	$\pi_1 = 222,318$
members'	–	–	$\pi_2 = 224.012$	–	$\pi_2 = 223.716$
payoff	$\pi_3 = 225.49$	$\pi_3 = 225.213$	$\pi_3 = 225.322$	$\pi_3 = 225.237$	$\pi_3 = 225.192$
Em. Red.	$Q = 57.647$	$Q = 58.289$	$Q = 57.95$	$Q = 59.072$	$Q = 58.663$

	6	7	8	9	10
Structure	$n_1 = 0$	$n_1 = 3$	$n_1 = 2$	$n_1 = 1$	$n_1 = 0$
	$n_2 = 2$	$n_2 = 0$	$n_2 = 1$	$n_2 = 2$	$n_2 = 3$
	$n_3 = 1$	$n_3 = 0$	$n_3 = 0$	$n_3 = 0$	$n_3 = 0$
Coalition	–	$\pi_1 = 222.222$	$\pi_1 = 222.181$	$\pi_1 = 222.226$	–
members'	$\pi_2 = 223.84$	–	$\pi_2 = 223.739$	$\pi_2 = 223.698$	$\pi_2 = 223.742$
payoff	$\pi_3 = 225.226$	–	–	–	–
Em. Red.	$Q = 58.289$	$Q = 60$	$Q = 59.518$	$Q = 59.072$	$Q = 58.663$

Price per TPP can be determined in two ways, it depends on market scope and if a finite or infinite number of players is presented. In case there is an infinite number, the market price can be supposed to be a fixed value p_∞, proposed by a certain consulting assembly, for example. When the emission trading market can be accessed by a finite number $n = \sum_{i=1}^{K} n_i$ of players, which belong to the formed agreement, the market price p per TPP unit is determined according to the economics optimality principle.

Optimal market price p lies on the intersection of marginal benefit of coalition and collective marginal abatement cost curves, *e.g.* [Hanley *et al.* 1997], [Copeland & Taylor 2003]. Thus to find price per TPP unit, it is necessary to solve the equation

$$\underbrace{MB(Q)\sum_{i=1}^{K} n_i}_{\text{marginal benefit of coalition}} = \underbrace{\sum_{i=1}^{K} n_i MC_i(q_i^S)}_{\text{marginal abatement costs of coalition}}. \tag{35}$$

Values $q_i = q_i^S$ are optimal individual abatements of signatories and $Q = Q_S + Q_F$ is abatement undertaken by all players, both signatories and free-riders. According to [Hanley *et al.* 1997], price can be determined as

$$p = MC_1(q_1^S) = MC_2(q_2^S) = \ldots = MC_K(q_K^S). \tag{36}$$

System of equations (35) delivers a solution which is equivalent to the solution of problem (24) from Section 3.3.3. Using results of Section 3.3.3, one can say that optimal abatement levels for individual signatories are

$$q_i^S = \frac{a\lambda_i(1-g)^2(\bar{1},\mathbf{n})}{(\bar{1},\mathbf{N}) + (1-g)^2(\bar{1},\mathbf{n})(\lambda,\mathbf{n})}, \quad i = 1,\ldots,K,$$

and the total abatement of signatories is

$$Q_S = \frac{a(1-g)^2(\bar{1},\mathbf{n})(\lambda,\mathbf{n})}{(\bar{1},\mathbf{N}) + (1-g)^2(\bar{1},\mathbf{n})(\lambda,\mathbf{n})},$$

where

$$g = \frac{(\lambda,\mathbf{N}-\mathbf{n})}{(\bar{1}+\lambda,\mathbf{N}) - (\lambda,\mathbf{n})}.$$

Free-riders adjust their abatement levels after having observed the choice of signatories, each free-rider maximizes its payoff non-cooperatively. Thus abatement reaction function of free-riders is

$$Q_F = g(a - Q_S).$$

According to (36) market price of pollution permit unit is

$$p = c_i q_i^S,$$

substituting expression of individual abatement q_i^S of a signatory delivers value

$$p = \frac{ab(1-g)^2(\bar{1},\mathbf{n})}{(\bar{1},\mathbf{N}) + (1-g)^2(\bar{1},\mathbf{n})(\lambda,\mathbf{n})}.$$

The IEA coalition is composed of $\sum_{i=1}^{K} n_i$ players. It means that total coalition abatement Q_S should be allocated among signatories

$$\sum_{i=1}^{K} q_i^{So} n_i = Q_S. \qquad (37)$$

Players will reduce their emissions up to the level where marginal costs of reduction are equal to the market price of permits (market is made up of coalition members). Players' emissions must be covered by equivalent amount of permits. Individual demand for pollution permits is $r_i = q_i^{So} - q_i^S$. If the player does not initially receive enough permits to cover its emissions, it must buy more permits from the market. If the player gets more permits than it needs to cover its emissions, it can sell the excess permits.

Individual cost of environmental protection is composed of net costs of reducing emissions and selling or buying emission permits. Costs of emission reductions are given by the area between the marginal abatement curve and the x-axis for the section between unregulated emission level (where $MC_i = 0$) and abatement level at which MC_i is equal to the market price for permits, p. The cost/benefit of buying/selling permits is simply the permit price multiplied by the quantity of permits bought/sold. Individual payoff of a signatory is

$$\pi_i(\mathbf{q}^S, \mathbf{q}^F) - pr_i,$$

where $(\mathbf{q}^S, \mathbf{q}^F)$ is a vector of optimal individual abatements q_i^S and q_i^F of signatories and free-riders respectively, p is a market price per TPP unit, and r_i is quantity of TPPs of signatory of type i.

The question of fair allocation of TPP is still opened and for this purpose we apply a side payments approach. According to (30), (34), the total surplus of coalition is

$$\Delta^S = \sum_{i=1}^{K} \Delta_i^S n_i,$$

where individual surplus Δ_i^S is from (31), and individual share of total surplus is

$$\delta_i^S = \frac{\lambda_i}{\sum_{l=1}^{K} n_l \lambda_l} \Delta^S.$$

Then fair allocation of pollution commitments according to agreement can be found from the equation

$$pr_i = \Delta_i^S - \delta_i^S,$$

where, as mentioned before, $r_i = q_i^{So} - q_i^S$. Hence

- market TPP price is

$$p = \frac{ab(1-g)^2(\bar{1}, \mathbf{n})}{(\bar{1}, \mathbf{N}) + (1-g)^2(\bar{1}, \mathbf{n})(\lambda, \mathbf{n})},$$

- signatory's commitment for pollution reduction, prescribed by IEA, is

$$q_i^{So} = q_i^S + \frac{1}{p}\left(\Delta_i^S - \delta_i^S\right).$$

  It is easy to check that condition (37) holds for such distribution. It is necessary to point out that in some particular cases values q_i^{So} can be non-positive, which implies that according to IEA some nations are allowed to increase their pollution to a certain level to improve global pollution situation. TPP amount for coalition members should be

$$r_i = \frac{1}{p}\left(\Delta_i^S - \delta_i^S\right),$$

- individual payoff of signatory is

$$\sigma_i^S = \pi_i(\mathbf{q}^{S\setminus i}, \mathbf{q}^{F\cup\{i\}}) + \delta_i^S,$$

  which is similar to (32).

Example 4.2.1

Let us turn to the case when three types of players are involved and reconsider the already familiar Example 4.1.3.

As was mentioned before, it means that we distinguish a group of industrialized countries (like USA, European Union, Japan, without pollution permits); a group of rapidly developing counties (like Russia, China), whose pollution permits stock is big enough but emission trading might be inefficient because pollution permits would be necessary for internal use to compensate extra emission discharge caused by industrial growth; and a group of agricultural countries with low abatement cost and large pollution permits stock.

TABLE 15 Example 4.2.1 Model Parameters

	Type 1	*Type* 2	*Type* 3
$b = 1$	$c_1 = 0.5$	$c_2 = 0.8$	$c_3 = 2$
$a = 100$	$N_1 = 5$	$N_2 = 10$	$N_3 = 3$

The simulation of stable coalition structure reveals diverse coalition structure (see Table 16). Now we shall accompany already given data of coalition structures, individual payoffs under side payments and global abatement, with agreement details, like amount of TPP units and market price per unit.

Real world situation related to emission trading among industrialized, rapidly developing, and agricultural countries reveals that emission trading might be inefficient for developing countries because pollution permits would be necessary for their internal use to compensate extra emission discharge caused by industrial growth. Investigations of this chapter have been carried out for static case only, thus to provide more detailed and specific analysis of possible scenarios, it is desirable that further research ought to concern dynamics of the process (see [de Zeeuw 2008], [Rubio & Ulph 2007], [Breton *et al.* 2008b], [Rubio & Ulph 2002]) possibly accompanied with introduced technological change, *e.g.* [Golombek & Hoel 2005], [Copeland & Taylor 2003], [Grübler & Gritsevskyi 2002].

TABLE 16   Example 4.2.1 Self-Enforcing Coalitions

	1	2	3	4	5
Structure	$n_1 = 0$ $n_2 = 0$ $n_3 = 3$	$n_1 = 1$ $n_2 = 0$ $n_3 = 2$	$n_1 = 0$ $n_2 = 1$ $n_3 = 2$	$n_1 = 2$ $n_2 = 0$ $n_3 = 1$	$n_1 = 1$ $n_2 = 1$ $n_3 = 1$
Coalition members' payoff	– – $\pi_3 = 225.49$	$\pi_1 = 222.479$ – $\pi_3 = 225.213$	– $\pi_2 = 224.012$ $\pi_3 = 225.322$	$\pi_1 = 222.193$ – $\pi_3 = 225.237$	$\pi_1 = 222,318$ $\pi_2 = 223.716$ $\pi_3 = 225.192$
Em. Red.	$Q = 57.647$	$Q = 58.289$	$Q = 57.95$	$Q = 59.072$	$Q = 58.663$
Abatement by IEA	– – $q_3^{So} = 1,569$	$q_1^{So} = 5.381$ – $q_3^{So} = 2.122$	– $q_2^{So} = 3.449$ $q_3^{So} = 1.846$	$q_1^{So} = 6.04$ – $q_3^{So} = 2.654$	$q_1^{So} = 5.708$ $q_2^{So} = 4.06$ $q_3^{So} = 2.39$
TPP	– – –	$q_1 = -1.036$ – $q_3 = 0.518$	– $q_2 = -0.518$ $q_3 = 0.259$	$q_1 = -0.509$ – $q_3 = 1.017$	$q_1 = -0.777$ $q_2 = 0.008$ $q_3 = 0.769$
Price	–	$p = 3.242$	$p = 3.173$	$p = 3.274$	$p = 3.242$

	6	7	8	9	10
Structure	$n_1 = 0$ $n_2 = 2$ $n_3 = 1$	$n_1 = 3$ $n_2 = 0$ $n_3 = 0$	$n_1 = 2$ $n_2 = 1$ $n_3 = 0$	$n_1 = 1$ $n_2 = 2$ $n_3 = 0$	$n_1 = 0$ $n_2 = 3$ $n_3 = 0$
Coalition members' payoff	– $\pi_2 = 223.84$ $\pi_3 = 225.226$	$\pi_1 = 222.222$ – –	$\pi_1 = 222.181$ $\pi_2 = 223.739$ –	$\pi_1 = 222.226$ $\pi_2 = 223.698$ –	– $\pi_2 = 223.742$ –
Em. Red.	$Q = 58.289$	$Q = 60$	$Q = 59.518$	$Q = 59.072$	$Q = 58.663$
Abatement by IEA	– $q_2^{So} = 3.754$ $q_3^{So} = 2.118$	$q_1^{So} = 6.667$ – –	$q_1^{So} = 6.353$ $q_2^{So} = 4.643$ –	$q_1^{So} = 6.03$ $q_2^{So} = 4.352$ –	– $q_2^{So} = 4.053$ –
TPP	– $q_2 = -0.257$ $q_3 = 0.514$	– – –	$q_1 = -0.256$ $q_2 = 0.512$ –	$q_1 = -0.519$ $q_2 = 0.26$ –	– – –
Price	$p = 3.209$	–	$p = 3.305$	$p = 3.274$	–

Part II

Dynamics of International Environmental Agreement

5 DYNAMIC GAME AND TIME-CONSISTENCY

IEAs open up possibilities to analyze prospects and features of long-term collaboration among countries. In light of dynamic framework, such familiar issues of IEA formation as

- promotion of agreement membership, [Hoel & Schneider 1997], [Barrett 1994b],
- influence of participants heterogeneity on abatement target settings, [Finus 2001], [Barrett 1997a], [McGinty 2006], [Eyckmans & Tulkens 2003],
- introduction of incentive mechanisms, [Hoel 1992], [Bauer 1992], [Carraro & Siniscalco 1993], [Barrett 1997a], [Botteon & Carraro 1997], [Barrett 1997b], [Le Breton & Soubeyran 1997], [Katsoulacos 1997], [Carraro & Siniscalco 1998], [Jeppensen & Andersen 1998], [Finus 2004], [Botteon & Carraro 1998], [Finus & Rundshagen 2000], [Bosello *et al.* 2003], [Carraro *et al.* 2003], [Kolstad 2003], [Altamirano-Cabrera & Finus 2006], [Eyckmans & Finus 2004], [Carraro & Marchiory 2004], [Weikard 2005], [Weikard *et al.* 2006], [Carraro *et al.* 2006],

receive new ways of interpretation and require additional solution methods. On top of that, such novel topics as dynamics of the accumulated pollution stock, optimal abatement dynamics, and IEA evolution emerge.

An understanding of the fact that multilateral agreements are targeted to stepwise emission reduction over a finite number of time periods motivates us to generalize the static model, presented in Section 3.1, to dynamic framework. We constitute a non-cooperative multistage game with a finite time horizon and pollution flow (given as a difference between pollution levels at the previous and current steps). To determine how to proceed with the abatement compliance and how the abatement compliance process affects players' incentives to change their signatory/free-rider status during the agreement life-cycle, we use the concept of *time-consistency*.

Independently from [de Zeeuw 2008], we address the problem of compliance with the emission reduction required by the agreement and for that purpose we construct *time-consistent reallocation of emission reduction* and compare the outcomes of the signatories and free-riders regarding such abatement dynamics. The time-consistent emission reduction scheme stipulates that the choice of abatement efforts during each time period is adjusted according to the emission reductions taken during the previous stages as well as dynamics of the pollution. Material of Part II is based on publications [Pavlova *et al.* 2008], [Dementieva *et al.* 2007], [Dementieva *et al.* 2008], [Pavlova 2008].

Dynamic Model of IEA

We suppose that a certain coalition S has been formed and its members have committed to reduce prespecified amount of emission during $[0, m]$ $(m > 1)$ time periods. Let us assume that the chosen abatement commitments $(\mathbf{q}^S, \mathbf{q}^F)$ are allocated over $[0, m]$ according to a certain scheme so that at each step $t = 0, \ldots, m-1$, each player of type i, $i = 1, \ldots, K$, reduces emissions by the amount $\left(\Delta q_i^S[t, t+1), \Delta q_i^F[t, t+1)\right)$ over the time period $[t, t+1)$. Then

$$\Delta Q_S[t, t+1) = \textstyle\sum_{i=1}^K n_i \Delta q_i^S[t, t+1),$$
$$\Delta Q_F[t, t+1) = \textstyle\sum_{i=1}^K (N_i - n_i) \Delta q_i^F[t, t+1)$$

are the aggregate abatement undertaken by the coalition S and the free-riders $\mathcal{N} \setminus S$ over the time period $[t, t+1)$.

Let us suppose that during $t = 0, \ldots, m$ the parameters b and c_i, which describe slopes of the marginal abatement benefit and cost functions (see (2) and (3)), are fixed. Pollution flow is specified by the parameter $a(t)$ $(a(0) = a)$ and associated with undertaken abatements during each time period $[t-1, t)$

$$a(t+1) = (1-\theta)a(t) - \Delta Q[t, t+1), \tag{38}$$

where $\Delta Q[t, t+1) = \Delta Q_S[t, t+1) + \Delta Q_F[t, t+1)$ and θ is the pollutant's natural rate of degradation $(\theta \in [0, 1))$. Let us define the aggregate emission reduction undertaken by player of type i during the first t steps by

$$q_i^S[0, t) = \sum_{l=0}^{t-1} \Delta q_i^S[l, l+1), \quad t = 1, \ldots, m;$$

$$q_i^F[0, t) = \sum_{l=0}^{t-1} \Delta q_i^F[l, l+1); \quad t = 1, \ldots, m;$$

the aggregate emission reduction, which player of type i is going to make during the rest $m - t$ steps (see Fig. 5) by

$$q_i^S[t, m] = \sum_{l=t}^{m-1} \Delta q_i^S[l, l+1),$$

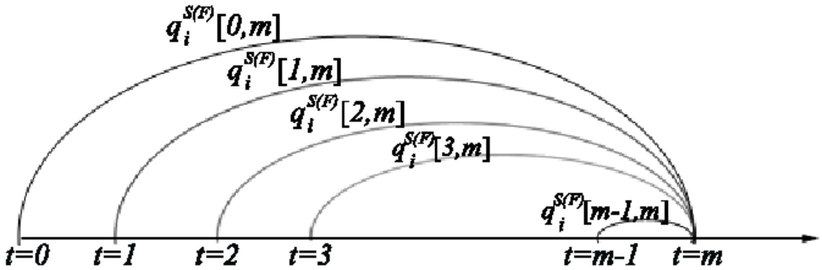


FIGURE 7 The abatement targets $\left(q_i^S[t,m], q_i^F[t,m]\right)$, $t = 0, \dots, m-1$.

$$q_i^F[t,m] = \sum_{l=t}^{m-1} \Delta q_i^F[l, l+1);$$

the aggregate emission reduction, which is to be performed by coalition S during the rest $m - t$ steps by

$$Q_S[t,m] = \sum_{i=1}^{K} n_i q_i^S[t,m];$$

the aggregate emission reduction, which is to be performed by free-riders during $m - t$ steps by

$$Q_F[t,m] = \sum_{i=1}^{K} (N_i - n_i) q_i^F[t,m];$$

and

$$Q[t,m] = Q_S[t,m] + Q_F[t,m].$$

Consider a current game $\Gamma_t\left(S, \mathbf{q}^S[0,t), \mathbf{q}^F[0,t)\right)$, where pollution flow[1] is indicated by $a(t)$ (38),

$$a(t) = (1-\theta)a - \left(\Delta Q_S[t-1,t) + \Delta Q_F[t-1,t)\right),$$

and the players' payoff over the remaining time period is assigned according to (1)

$$\pi_i(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) = \rho^t \left[\frac{b}{N}\left(a(t)(Q[t,m] - \frac{1}{2}(Q^2[t,m]\right) - \frac{1}{2} c_i (q_i[1,m])^2\right]. \tag{39}$$

Here θ is the pollutant's natural rate of degradation ($\theta \in [0,1)$) and ρ is discount factor ($\rho \in (0,1]$).

We assume that the chosen abatement commitments, determined by the Stackelberg equilibrium $(\mathbf{q}^S, \mathbf{q}^F)$ in the game $\Gamma_0(S)$ (see Lemma 3.3.1). A stepwise realization of a solution of the game $\Gamma_0(S)$ during the given time period $[0,m]$ can cause loss of optimality of the solution in the current game. In such a case the solution is said to be time-inconsistent, [Dementieva *et al.* 2007]. Absence

1 We point out that the static game $\Gamma_0(S) = \Gamma_0(S,0,0)$ when $t = 0$.

of time-consistency involves the possibility that the previous "optimal" decision is abandoned at some current moment of time, thereby making the problem for seeking an optimal control meaningless, [Strotz 1955].

Introduced in [Petrosjan 1977], the term of *time-consistency* of an optimality principle, is of high importance while analyzing dynamic games. It means that any segment of the optimal trajectory determines optimal motion with respect to the relevant initial states of the process. This property holds for the overwhelming majority of classical optimal control problems and follows from the Bellman optimality principle, [Bellman 2003]. Approaches to design a mechanism for allocation over time of players' payoffs, which guarantees time-consistency of solutions in cooperative and noncooperative games, were discussed in [Petrosjan & Zaccour 2003], [Zakharov 1988], [Dementieva 2004], [Petrosjan & Danilov 1979], [Petrosjan & Zaccour 2000]. Basing on such approaches, let us introduce a definition of a time-consistent scheme of stepwise emission reduction in the game $\Gamma_0(S)$.

Definition 5.0.1 *A scheme of stepwise emission reduction* $\{\Delta q_i^S[t, t+1), \Delta q_i^F[t, t+1)\}_{t=0}^{m-1}$ *of players of type* $i = 1, \dots, K$ *is called time-consistent, if the correspondent abatements* $\left(\mathbf{q}^S[t, m], \mathbf{q}^F[t, m]\right)$, $t = 0, \dots, m-1$, *constitute Stackelberg equilibrium in the reduced game* $\Gamma_t\left(S, \mathbf{q}^S[0, t), \mathbf{q}^F[0, t)\right)$.

Since stepwise emission reduction affects the pollution flow $a(t)$, it can likely motivate players to reconsider their signatory/free-rider status at a certain moment t and put stability of the agreement S in jeopardy. Withdrawal of some nations from the agreement and accessing of others would cause structural change of the coalition and sequential switch to another abatement goal. We say that a self-enforcing coalition is *time-consistent* if none of the signatories/free-riders has incentive to change its status and leave/access the coalition at any moment of time. In other terms it implies that if at the initial moment $t = 0$ in the game $\Gamma_0(S)$ players form a coalition S according to the self-enforcing optimality principle (see Definition 3.3.1) then at each current moment t the formed coalition must remain stable, *i.e.* satisfy the conditions of internal and external stability, (26) and (27), required by self-enforcement, in the game $\Gamma_t(S, \mathbf{q}^S[0, t), \mathbf{q}^F[0, t))$, $t = 0, \dots, m-1$.

Given that the pollution evolves according to the formula (38), we should modify the conditions of internal/external stability and explore players' motivation to reconsider their signatory/free-rider status. We assume that having accomplished its obligations for moment t, a signatory considers withdrawing from the coalition S if its payoff as a signatory of S over the time period $[t, m]$ is smaller than its payoff as a free-rider from set the $F \cup \{i\}$; and vice versa, a free-rider considers accessing the coalition S if its payoff as a free-rider from the set F over the time period $[t, m]$ is smaller than its payoff as a signatory of the coalition $S \cup \{i\}$.

Definition 5.0.2 *A self-enforcing coalition* S*, characterized by* $\mathbf{n} = (n_1, \dots, n_K)$ *of players of* K *types, is time-consistent under a given abatement scheme, if for every time* t, $t = 0, \dots, m-1$, *the following conditions hold simultaneously*

1) internal time-consistency

$$\pi_i(\mathbf{q}^S[t,m],\mathbf{q}^F[t,m]) \geq \pi_i(\mathbf{q}^{S\setminus\{i\}}[t,m],\mathbf{q}^{F\cup\{i\}}[t,m]), \quad \forall i \in S, \tag{40}$$

where $(\mathbf{q}^S[t,m],\mathbf{q}^F[t,m])$ *is the restriction to the period* $[t,m]$ *of an optimal solution of the game* $\Gamma_0(S)$ *and* $(\mathbf{q}^{S\setminus\{i\}}[t,m],\mathbf{q}^{F\cup\{i\}}[t,m])$ *is the restriction to the period* $[t,m]$ *of an optimal solution of the game* $\Gamma_0(S\setminus\{i\})$*,*

2) external time-consistency

$$\pi_i(\mathbf{q}^S[t,m],\mathbf{q}^F[t,m]) \geq \pi_i(\mathbf{q}^{S\cup\{i\}}[t,m],\mathbf{q}^{F\setminus\{i\}}[t,m]), \quad \forall i \in F, \tag{41}$$

where $(\mathbf{q}^{S\cup\{i\}}[t,m],\mathbf{q}^{F\setminus\{i\}}[t,m])$ *is restriction over the period* $[t,m]$ *of an optimal solution of the game* $\Gamma_0(S\cup\{i\})$*.*

6 PROPERTY OF TIME-CONSISTENCY IN MULTISTAGE DYNAMICS

6.1 Time-Consistent Emission Reduction Scheme

Construct a time-consistent scheme of stepwise emission reduction in the game $\Gamma_0(S)$. Defining approachs to reallocating abatement commitments over accounting periods requires understanding of players intentions. Indeed, what would be a reasonable perspective to judge of nations' realistic behavior? Imagine N nations have just made their decisions wether or not to join an IEA and according to these decisions each of them has an optimal abating target.

In other words, each player has got a certain set of cards, which can be disposed during the game, and knows that his payoff will depend on the actions of others, and that every card he plays will effect the following sequence of the game.

Hence after the players have indicated their equilibrium abatement efforts, they split their commitments into two parts, the former part specifies abatement over $[0, 1)$ and the latter one corresponds to the rest of the game. Choosing strategies $\left(\Delta q_i^S[0,1), \Delta q_i^F[0,1)\right)$ for the first period, the players anticipate that during $[1, m]$ they will have to couple with the remaining part of the abatement commitments $\left(q_i^S[1,m], q_i^F[1,m]\right)$ upon the pollution level $a(1)$, which is $a(1) = (1-\theta)a - \Delta Q_S[0,1) - \Delta Q_F[0,1)$.

We are going to construct a stepwise emission reduction scheme, that is time-consistent within the considered model, so that the choice of abatement efforts during each time period $[t, t+1)$ is adjusted according to the emission reduction, undertaken during the previous stage, and dynamics of the pollutant. Let us consider the current game $\Gamma_t\left(S, \mathbf{q}^S[0,t), \mathbf{q}^F[0,t)\right)$, where players' payoff is assigned according to (39) and the parameter $a(t)$, which indicates the current pollution flow, determined by (38).

In line with Definition 5.0.1, time-consistency of a scheme means that emission reduction $\left(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]\right)$ constructs Stackelberg equilibrium in the current game $\Gamma_t\left(S, \mathbf{q}^S[0,t), \mathbf{q}^F[0,t)\right)$. We introduce the following notations:

$$\tilde{\mu}_i = \frac{\lambda_i(1-\rho)}{(1+\rho)N + (1-\rho)(\lambda, \mathbf{N}-\mathbf{n})},$$
$$\mu = (\tilde{\mu}, \mathbf{N}-\mathbf{n}),$$
$$\eta = \frac{N\rho}{(1+\rho)N + (1-\rho)(\lambda, \mathbf{N}-\mathbf{n})}.$$

We remind that here $\rho \in (0,1]$ is the discount factor and θ is the pollutant's natural rate of degradation.

Theorem 6.1.1 *Consider a coalition S and the corresponding Stackelberg equilibrium $\left(\mathbf{q}^S, \mathbf{q}^F\right)$ of the game $\Gamma_0(S)$. The scheme of stepwise abatement*

$$\begin{aligned}\Delta q_i^S[t,t+1) &= \frac{N\rho}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2 + N(1+\rho)} q_i^S[t,m] \\ &+ \frac{\lambda_i(1-\rho)(1-\mu)(\bar{1},\mathbf{n})}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2 + N(1+\rho)} (((1-\theta)(1-\mu) \\ &\quad + \theta(1-2\mu))a(t) - \eta Q_F[t,m]),\end{aligned} \tag{42}$$

$$\begin{aligned}\Delta q_i^F[t,t+1) &= \tilde{\mu}_i\left((1-\theta)a(t) - \Delta Q_S[t,t+1)\right) \\ &+ \frac{\lambda_i\theta}{(1+\rho)N + (1-\rho)(\lambda,\mathbf{N}-\mathbf{n})} a + \eta q_i^F[t,m],\end{aligned}$$

is time-consistent, $t = 0,\ldots,m-1$, $i = 1,\ldots,K$.

P r o o f. The proof of the theorem is based on the principle of mathematical induction. Consider the time $t = 1$. Having determined abatement commitments $\left(\mathbf{q}^S, \mathbf{q}^F\right)$ in the game $\Gamma_0(S)$, the players split these commitments into two parts. The first part $\left(\Delta q_i^S[0,1), \Delta q_i^F[0,1)\right)$ specifies abatement during $[0,1)$ and the remaining part of the commitment $q_i^{S(F)}[1,m]$ is expected to be reduced by the end of the game. Choosing strategies $\left(\Delta q_i^{S(F)}[0,1), \Delta q_i^F[0,1)\right)$ for the first period, players anticipate that during $[1,m]$ they will have to deal with the remaining abatement commitments $\left(q_i^S[1,m], q_i^F[1,m]\right)$ while the pollution stock is given by

$$a(1) = (1-\theta)a - \Delta Q_S[0,1) - \Delta Q_F[0,1).$$

As before, we consider a two-level game, where the coalition S is the leader and the free-riders in the set F are the followers. Suppose the players in the coalition S decide about certain feasible strategies $\Delta q_i^S[0,1)$, $i = 1,\ldots,K$.

Having such information, free-riders should react rationally (choosing Nash strategies). In pursuit of the most beneficial situation for them in the reduced game $\Gamma_t\left(S, \mathbf{q}^S[0,t), \mathbf{q}^F[0,t)\right)$ during $[1,m]$, they adjust $\Delta q_i^F[0,1)$ by solving the following problem

$$\max_{q_i^F} \phi_i^F(1), \qquad i \in F \tag{43}$$

where

$$\begin{aligned}
\phi_i^F(1) &= \pi_i(\Delta\mathbf{q}^S[0,1), \Delta\mathbf{q}^F[0,1)) + \rho\pi_i(\mathbf{q}^S[1,m], \mathbf{q}^F[1,m]) \\
&= \frac{b}{N}\left(a(\Delta Q_S[0,1) + \Delta Q_F[0,1)) - \frac{1}{2}(\Delta Q_S[0,1) + \Delta Q_F[0,1))^2\right) \\
&- \frac{1}{2}c_i\left(\Delta q_i^F[0,1)\right)^2 \\
&+ \rho\frac{b}{N}\left(a(1)(Q_S[1,m] + Q_F[1,m]) - \frac{1}{2}(Q_S[1,m] + Q_F[1,m])^2\right) \\
&- \rho\frac{1}{2}c_i\left(q_i^F[1,m]\right)^2.
\end{aligned}$$

The signatories solve the maximization problem regarding the interests of the whole coalition S, given

$$\max_{q_i^S}\sum_{i=1}^{K} n_i\phi_i^S(1), \tag{44}$$

where

$$\begin{aligned}
\sum_{i=1} n_i\phi_i^S(1) &= \sum_{i=1}^{K} n_i\left(\pi_i(\Delta\mathbf{q}^S[0,1), \Delta\mathbf{q}^F[0,1)) + \pi_i(\mathbf{q}^S[1,m], \mathbf{q}^F[1,m])\right) \\
&= \frac{b}{N}(1,\mathbf{n})\left(a(\Delta Q_S[0,1) + \Delta Q_F[0,1)) - \frac{1}{2}(\Delta Q_S[0,1) + \Delta Q_F[0,1))^2\right) \\
&- \frac{1}{2}\sum_{i=1}^{K} n_i c_i\left(\Delta q_i^S[0,1)\right)^2 \\
&+ \rho\frac{b}{N}(1,\mathbf{n})\left(a(1)(Q_S[1,m] + Q_F[1,m]) - \frac{1}{2}(Q_S[1,m] + Q_F[1,m])^2\right) \\
&- \frac{1}{2}\rho\sum_{i=1}^{K} n_i c_i\left(q_i^S[1,m]\right)^2.
\end{aligned}$$

Let us define the followers' reaction functions for any choice of the leader's strategy. To identify $\Delta q_i^F[0,1)$ and $q_i^F[1,m]$, $i = 1,\ldots,K$, and thus $\Delta Q_F[0,1)$ and $Q_F[1,m]$, we solve maximization problem (43), assuming the leader's choice of the strategies $q_i^S[1,m]$ is known. Taking into account that $q_i^F = \Delta q_i^F[0,1) + q_i^F[1,m]$, the first order condition

$$\frac{\partial\phi_i^F(1)}{\partial\Delta q_i^F[0,1)} = 0, \qquad i = 1,\ldots,K,$$

leads to

$$\begin{aligned}
&\frac{b}{N}(a - \Delta Q[0,1)) + \rho\frac{b}{N}(-a(1-\theta) + \Delta Q[0,1)) - c_i\Delta q_i^F[0,1) \\
&+ c_i\rho(q_i^F - \Delta q_i^F[0,1)) = 0,
\end{aligned}$$

which is equivalent to

$$\frac{b}{N}(1-\rho)(a(1-\theta) - \Delta Q[0,1)) + \frac{b}{N}\theta a - (1+\rho)c_i\Delta q_i^F[0,1) + c_i\rho q_i^F = 0$$

and

$$\frac{b}{N}(1-\rho)(a(1-\theta)-\Delta Q^S[0,1))+\frac{b}{N}\theta a+\rho c_i q_i^F-\frac{b}{N}(1-\rho)\Delta Q^F[0,1)$$
$$=(1+\rho)c_i\Delta q_i^F[0,1).$$

Since $c_i q_i^F = c_j q_j^F$ ($i \neq j$) holds (see, for instance, Section 3.2), then it also holds that $\Delta q_i^F[0,1) = \Delta q_j^F[0,1)$. Consequently, the first order condition turns out to be as follows

$$\frac{b}{N}(1-\rho)((1-\theta)a-\Delta Q^S[0,1))+\frac{b}{N}\theta a+\rho c_i q_i^F$$
$$=(1+\rho)c_i\Delta q_i^F[0,1)+(1-\rho)\frac{c_i}{N}(\lambda,\mathbf{N}-\mathbf{n})\Delta q_i^F[0,1).$$

Thus, we obtain the following solution

$$\Delta q_i^F[0,1)=\frac{\lambda_i(1-\rho)((1-\theta)a-\Delta Q_S[0,1))+\lambda_i\theta a}{(1+\rho)N+(1-\rho)(\lambda,\mathbf{N}-\mathbf{n})} \tag{45}$$

$$+\frac{N\rho q_i^F}{(1+\rho)N+(1-\rho)(\lambda,\mathbf{N}-\mathbf{n})},$$

where $i=1,\dots,K$. Since

$$\frac{\partial^2\phi_i^F(1)}{\partial^2\Delta q_i^F[0,1)}=-2c_i(1+\rho)-b/N(1-\rho)-(1-\rho)(\lambda,\mathbf{N}-\mathbf{n})c_i/N<0$$

for every $i=1,\dots,K$, then (45) is a solution of the maximization problem (43). Thus during period $[0,1)$ free-riders reduce their emissions according to (45), and their aggregate abatement during $[0,1)$ is

$$\Delta Q_F[0,1)=\frac{(1-\rho)(\lambda,\mathbf{N}-\mathbf{n})((1-\theta)a-\Delta Q_S[0,1))+(\lambda,\mathbf{N}-\mathbf{n})\theta a}{(1+\rho)N+(1-\rho)(\lambda,\mathbf{N}-\mathbf{n})}$$
$$+\frac{N\rho Q_F}{(1+\rho)N+(1-\rho)(\lambda,\mathbf{N}-\mathbf{n})},$$

where Q_F is given in (23), (25).

In order to determine leader's strategies $\Delta q_i^S[0,1)$, $i=1,\dots,K$ we need to solve the maximization problem (44), taking the rational reaction functions (45) of the followers into account. The first order conditions

$$\frac{\partial\sum_{i=1}n_i\phi_i^S(1)}{\partial\Delta q_i^S[0,1)}=0,\qquad i=1,\dots,K,$$

delivers

$$
\begin{aligned}
&\frac{b}{N}(1,\mathbf{n})n_i(1-\mu)\left[(1-\mu)(1-\theta)a-\frac{\mu}{1-\rho}\theta a-\eta Q_F\right]\\
&-\rho\frac{b}{N}(1,\mathbf{n})n_i(1-\mu)\left[(1-\mu)(1-\theta)a-\frac{\mu}{1-\rho}\theta a-\eta Q_F\right]\\
&+\frac{b}{N}(1,\mathbf{n})n_i(1-\mu)^2\theta a+\rho c_i n_i q_i^S\\
&-\frac{b}{N}(1,\mathbf{n})(1-\mu)^2(1-\rho)n_i\Delta Q_S-(1+\rho)n_i c_i\Delta q_i^S\\
&=(1-\rho)\frac{b}{N}(1,\mathbf{n})(1-\mu)\left[(1-\mu)(1-\theta)a-\frac{\mu}{1-\rho}\theta a-\eta Q_F\right]\\
&+\frac{b}{N}(1,\mathbf{n})n_i(1-\mu)^2\theta a+\rho n_i c_i q_i^S\\
&-\frac{c_i}{N}(1,\mathbf{n})n_i(1-\mu)^2(1-\rho)(\lambda,\mathbf{n})\Delta q_i^S-(1+\rho)n_i c_i\Delta q_i^S=0.
\end{aligned}
$$

Consequently, the equation can be represented in the following manner

$$
\begin{aligned}
&\lambda_i(1-\rho)(\bar{1},\mathbf{n})(1-\mu)\left[(1-\mu)(1-\theta)a-\frac{\mu}{1-\rho}\theta a-\eta Q_F\right]+\rho c_i q_i^S\\
&+\lambda_i(\bar{1},\mathbf{n})(1-\mu)^2\theta a=\Delta q_i^S((\bar{1},\mathbf{n})(1-\mu)^2(1-\rho)(\lambda,\mathbf{n})+N(1+\rho)).
\end{aligned}
$$

Hence, the maximization problem (44) has the following solution

$$
\begin{aligned}
\Delta q_i^S[0,1)=&\frac{N\rho}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}q_i^S \qquad (46)\\
&+\frac{\lambda_i(1-\rho)(1-\mu)^2(\bar{1},\mathbf{n})(1-\theta)}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}a\\
&+\frac{\lambda_i(\bar{1},\mathbf{n})(1-\mu)(1-2\mu)\theta}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}a\\
&-\frac{\lambda_i(\bar{1},\mathbf{n})(1-\mu)(1-\rho)\eta}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}Q_F
\end{aligned}
$$

for $i=1,\ldots,K$. Since the second partial derivatives are negative, then (46) is a solution of maximization problem (44) and describes individual abating efforts of signatories during $t\in[0,1)$. All the signatories together will abate

$$
\begin{aligned}
\Delta Q_S[0,1)=&\frac{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\mu)^2(1-\rho)(1-\theta)}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}a\\
&+\frac{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-2\mu)\theta}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}a\\
&-\frac{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\mu)\eta}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}Q_F\\
&+\frac{\rho N}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}Q_S,
\end{aligned}
$$

where Q_S is given by (25).

Assume that for every $t = 1, \dots, \tau \le m-1$ and $i = 1, \dots, K$

$$\begin{aligned}
\Delta q_i^S[t,t+1) = & \frac{N\rho}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2 + N(1+\rho)} q_i^S[t,m] \\
& + \frac{\lambda_i(1-\rho)(1-\mu)^2(\bar{1},\mathbf{n})(1-\theta)}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2 + N(1+\rho)} a(t) \\
& + \frac{\lambda_i(\bar{1},\mathbf{n})(1-\mu)(1-2\mu)\theta}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2 + N(1+\rho)} a(t) \\
& - \frac{\lambda_i(\bar{1},\mathbf{n})(1-\mu)(1-\rho)\eta}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2 + N(1+\rho)} Q_F[t,m], \\
\Delta q_i^F[t,t+1) = & \frac{\lambda_i(1-\rho)}{(1+\rho)N + (1-\rho)(\lambda,\mathbf{N}-\mathbf{n})} \left((1-\theta)a(t-1) - \Delta Q_S[t,t+1)\right) \\
& + \frac{\lambda_i\theta}{(1+\rho)N + (1-\rho)(\lambda,\mathbf{N}-\mathbf{n})} a(t) \\
& + \frac{N\rho}{(1+\rho)N + (1-\rho)(\lambda,\mathbf{N}-\mathbf{n})} q_i^F[t,m].
\end{aligned}$$

We now show that for $t = \tau + 1$ it also holds. Following a similar approach, we determine abatement levels $\Delta q_i^{S(F)}[\tau,\tau+1)$, so that the restricted solution $\left(q_i^S[\tau+1,m], q_i^F[\tau+1,m]\right)$ constitutes Stackelberg equilibrium in the current game $\Gamma_\tau\left(S, \mathbf{q}^S[0,\tau), \mathbf{q}^F[0,\tau)\right)$. The pollution level at the time $t = \tau$ and $t = \tau + 1$ is given by

$$a(\tau) = (1-\theta)a(\tau-1) - \Delta Q[\tau-1,\tau),$$

and

$$a(\tau+1) = (1-\theta)a(\tau) - \Delta Q[\tau,\tau+1).$$

Each free-rider of type i maximizes its individual payoff gained during the periods $[\tau,\tau+1)$ and $[\tau+1,m]$, taking the abatement choice of the leader as given

$$\max_{\Delta q_i^F[\tau,\tau+1)} \phi_i^F(\tau+1). \tag{47}$$

These payoffs are expressed by

$$\begin{aligned}
\phi_i^F(\tau+1) = & \rho^\tau \frac{b}{N}\Bigg[a(\tau)(\Delta Q_S[\tau,\tau+1) + \Delta Q_F[\tau,\tau+1)) \\
& - \frac{1}{2}(\Delta Q_S[\tau,\tau+1) + \Delta Q_F[\tau,\tau+1))^2\Bigg] - \rho^\tau \frac{1}{2} c_i \left(\Delta q_i^F[\tau,\tau+1)\right)^2 \\
& + \rho^{\tau+1}\frac{b}{N}\Bigg[a(\tau+1)(Q_S[\tau+1,m] + Q_F[\tau+1,m]) \\
& - \frac{1}{2}(Q_S[\tau+1,m] + Q_F[\tau+1,m])^2 \Bigg] - \rho^{\tau+1}\,\frac{1}{2} c_i \left(q_i^F[\tau+1,m]\right)^2.
\end{aligned}$$

A solution of the maximization problem (47) is

$$\Delta q_i^F[\tau,\tau+1) = \frac{N\rho}{(1+\rho)N + (1-\rho)(\lambda,\mathbf{N}-\mathbf{n})} q_i^F[\tau,m] \tag{48}$$

$$+\frac{\lambda_i\theta}{(1+\rho)N+(1-\rho)(\lambda,\mathbf{N}-\mathbf{n})}a(\tau)$$
$$+\frac{\lambda_i(1-\rho)(1-\theta)}{(1+\rho)N+(1-\rho)(\lambda,\mathbf{N}-\mathbf{n})}a(\tau)$$
$$-\frac{\lambda_i(1-\rho)}{(1+\rho)N+(1-\rho)(\lambda,\mathbf{N}-\mathbf{n})}\Delta Q_S[\tau,\tau+1).$$

Thus during $[\tau,\tau+1)$ free-riders will undertake emission reduction according to (48). To determine the strategies $\Delta q_i^S[\tau,\tau+1)$, $i=1,\dots,K$, the signatories maximize coalitional net benefit using information about the free-riders' rational reaction, leading the maximization problem

$$\max_{\Delta q_i^S[\tau,\tau+1)}\sum_{i=1}^{K}n_i\phi_i^S(\tau+1)$$

where

$$\sum_{i=1}^{K}n_i\phi_i^S(\tau+1) \tag{49}$$

$$=\frac{b}{N}(1,\mathbf{n})\left[a(\tau)(\Delta Q_S[\tau,\tau+1)+\Delta Q_F[\tau,\tau+1))-\frac{1}{2}(\Delta Q_S[\tau,\tau+1)+\Delta Q_F[\tau,\tau+1))^2\right]$$
$$-\frac{1}{2}\sum_{i=1}^{K}n_ic_i\left(\Delta q_i^S[\tau,\tau+1)\right)^2$$
$$+\frac{b}{N}(1,\mathbf{n})\left[a(\tau+1)(Q_S[\tau+1,m]+Q_F[\tau+1,m])-\frac{1}{2}(Q_S[\tau+1,m]+Q_F[\tau+1,m])^2\right]$$
$$-\frac{1}{2}\sum_{i=1}^{K}n_ic_i\left(q_i^S[\tau+1,m]\right)^2.$$

A solution of the maximization problem (49) is

$$\Delta q_i^S[\tau,\tau+1)=\frac{N\rho}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}q_i^S[\tau,m] \tag{50}$$

$$+\frac{\lambda_i(1-\rho)(1-\mu)^2(\bar{1},\mathbf{n})(1-\theta)}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}a(\tau)$$
$$+\frac{\lambda_i(\bar{1},\mathbf{n})(1-\mu)(1-2\mu)\theta}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}a(\tau)$$
$$-\frac{\lambda_i(\bar{1},\mathbf{n})(1-\mu)(1-\rho)\eta}{(\lambda,\mathbf{n})(\bar{1},\mathbf{n})(1-\rho)(1-\mu)^2+N(1+\rho)}Q_F[\tau,m].$$

Hence, abatement undertaken by a signatory of type i during $[\tau,\tau+1)$ is given by (50). According to the principle of mathematical induction we have shown that the emission reduction scheme (51) is time-consistent in the game $\Gamma_0(S)$ by definition. ■

Corollary 6.1.1 *Let a coalition S and the correspondent Stackelberg equilibrium* $\left(\mathbf{q}^S, \mathbf{q}^F\right)$ *be formed in the game* $\Gamma_0(S)$*. We set discount factor* $\rho = 1$ *and natural degradation rate* $\theta = 0$.

The time-consistent scheme of stepwise abatement reduces to

$$\Delta q_i^{S(F)}[t, t+1] = \frac{1}{2^t} q_i^{S(F)}, \qquad \text{where } t = 0, \dots, m-2, \tag{51}$$

$$\Delta q_i^{S(F)}[m-1, m] = \frac{1}{2^{m-1}} q_i^{S(F)}.$$

Notice that the time-consistent scheme (51) is described as a geometric progression with 0.5 as a common ratio and $\left(\mathbf{q}^S, \mathbf{q}^F\right)$ as an initial element. According to such a scheme, large emission reduction should be undertaken during first stages and the following sequence of abating efforts will monotonically decrease. Additionally, from the representation of the time-consistent scheme in Corollary 6.1.1 it follows that the determined Stackelberg equilibrium in the current game $\Gamma_t\left(S, \mathbf{q}^S[0,t), \mathbf{q}^F[0,t)\right)$ coincides with Nash equilibrium, because free-riders and signatories choose their strategies simultaneously and independently of each other.

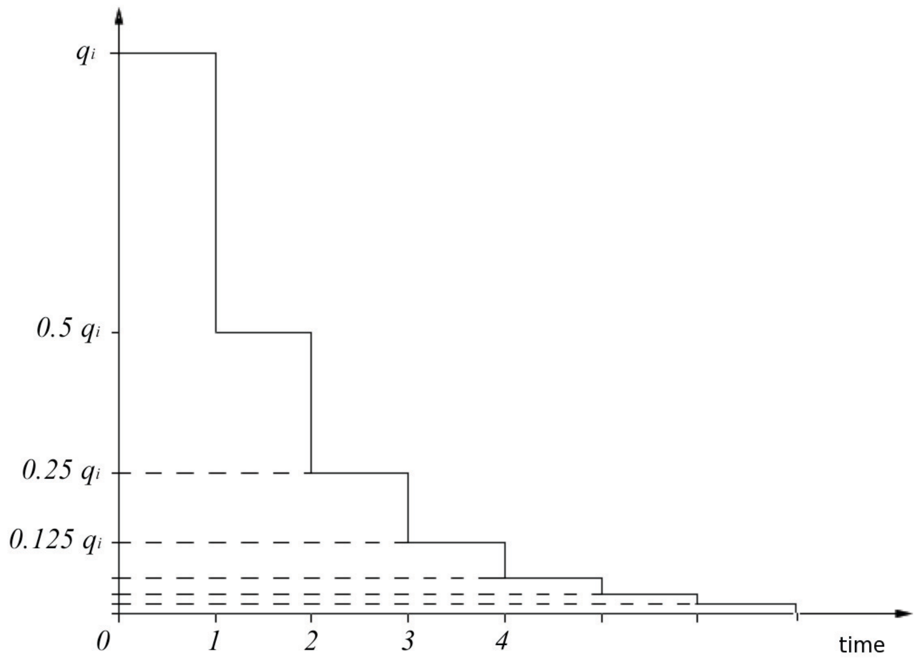


FIGURE 8 The time-consistent abatement scheme (51), $\rho = 1$, $\theta = 0$.

Example 6.1.1

Consider the numerical example with the model settings given in Table 17.

TABLE 17 Example 6.1.1 Model Parameters

	Type 1	*Type* 2
$b = 1$	$c_1 = 0.5$	$c_2 = 0.8$
$a = 100$	$N_1 = 5$	$N_2 = 10$

In this example there are two self-enforcing coalitions, Table 18. Let us check the IEA S_2 for time-consistency and explore properties of its emission reduction trajectory.

TABLE 18 Example 6.1.1 Self-Enforcing Coalitions

Coalition structure	S_1 $(n_1 = 0, n_2 = 3)$	S_2 $(n_1 = 1, n_2 = 2)$
Signatory's payoff	— $\pi_2^S = 275.986$	$\pi_1^S = 272.226$ $\pi_2^S = 276.779$
Free-rider's payoffs	$\pi_1^F = 276.726$ $\pi_2^F = 279.223$	$\pi_1^F = 277.838$ $\pi_2^F = 280.286$
Signatories' abatement	— $q_2^S = 4.301$	$q_1^S = 6.969$ $q_2^S = 4.355$
Free-riders' abatement	$q_1^F = 5.161$ $q_2^F = 3.22$	$q_1^F = 5.11$ $q_2^F = 3.194$

Suppose that the abatement commitments must be fulfilled over m steps according to the time-consistent abatement scheme (42), when discounting and pollutant degradation rates are $\rho = 0.95$ and $\theta = 0.05$. Let us assume potential free-riding from membership and have a look what may occur at the step $t = 1$:

- signatories' abatement during $[0, 1)$ is $\Delta q_1^S[0, 1) = 5.395$, $\Delta q_2^S[0, 1) = 3.372$,
- free-riders' abatement during $[0, 1)$ is $\Delta q_1^F[0, 1) = 3.022$, $\Delta q_2^F = 1.889$,
- signatories' abatement during $[1, m]$ is $\Delta q_1^S[1, m] = 1.574$, $\Delta q_2^S[1, m] = 0.983$,
- free-riders' abatement during $[1, m]$ is $\Delta q_1^F[1, m] = 2.069$, $\Delta q_2^F[1, m] = 1.293$. It is important to point out the following interesting observation: according to the scheme (42), larger part of emission should be reduced on the first abating interval!
- Abatement of possible free-riders from the agreement is

$$\Delta q_1^{F\cup 1}[1, m] = 2.211, \Delta q_2^{F\cup 2}[1, m] = 1.383,$$

- abatement of possible new members of the agreement is

$$\Delta q_1^{S\cup 1}[1,m] = 2.055, \Delta q_2^{S\cup 2}[1,m] = 1.255,$$

- additionally, $\Delta Q[1,m] = 39.335$, $Q[1,m] = 22.164$, $Q^{-1}[1,m] = 23.327$, $Q^{-2}[1,m] = 22.988$.

To check time-consistency of the self-enforcing coalition S_2 we substitute thus found data into inequalities (40)–(41):
1) internal stability fails

$$\pi_1(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) = 72.645 \leq \pi_1(\mathbf{q}^{S\setminus 1}[t,m], \mathbf{q}^{F\cup 1}[t,m]) = 74.98,$$

$$\pi_2(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) = 73.022 \leq \pi_2(\mathbf{q}^{S\setminus 2}[t,m], \mathbf{q}^{F\cup 2}[t,m]) = 74.132,$$

2) external stability holds

$$\pi_1(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) = 72.193 \geq \pi_1(\mathbf{q}^{S\cup 1}[t,m], \mathbf{q}^{F\setminus 1}[t,m]) = 69.452,$$

$$\pi_2(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) = 72.845 \geq \pi_2(\mathbf{q}^{S\cup 2}[t,m], \mathbf{q}^{F\setminus 2}[t,m]) = 70.691.$$

Hence we can conclude that after the abatement process has been started the coalition may no longer satisfy its initial stability criterion and the players, which used to be signatories, have incentives to leave the coalition. Similar conclusions can be derived when considering potential free-riding from the compliance. Loss of time-consistency can lead to the renegotiation of the agreement structure and targets to regulate coalition instability.

In the following Sections 6.2 and 6.3 we analyze a special case of the time-consistent emission reduction scheme when discounting and pollutant degradation rates are equal to zero.

6.2 Time-Consistency of Internally Stable Agreement

Given the Stackelberg solution in the game $\Gamma_0(S)$ and assuming that the coalition S remains the same over the time period $[0,m]$, we produced in Section 6.1 a time-consistent scheme that allows players to fulfil their emission reduction targets gradually.

To analyze the possibility of withdrawal of some nations from the agreement S and the joining of the new nations, we have given a notion of time-consistency of a self-enforcing coalition and applied it to the current game $\Gamma_t(S, \mathbf{q}^S[0,t), \mathbf{q}^F[0,t))$. According to Definition 5.0.2, an internally stable coalition S, characterized by vector $\mathbf{n} = (n_1, \dots, n_K)$ of players of K types, is time-consistent under a given abatement scheme, if for every time t, $t = 0, \dots, m-1$, the following conditions hold simultaneously:

$$\pi_i(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) \geq \pi_i(\mathbf{q}^{S\setminus\{i\}}[t,m], \mathbf{q}^{F\cup\{i\}}[t,m]), \quad \forall i \in S.$$

Here $(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m])$ is the restriction to the period $[t,m]$ of an optimal solution of the game $\Gamma_0(S)$, and $(\mathbf{q}^{S\setminus\{i\}}[t,m], \mathbf{q}^{F\cup\{i\}}[t,m])$ is the restriction to the period $[t,m]$ of an optimal solution of the game $\Gamma_0(S\setminus\{i\})$. We compare the free-riders' outcomes with the outcomes of signatories, supposing that they are calculated under abatement strategies determined by the restriction to the period $[t,m]$ of the Stackelberg solutions, $t=1,\dots,m-1$. Since parameters c_i, b are constant, for each signatory of type i, $i\in S$, the functions in the internal time-consistency conditions are

$$\pi_i(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) = \frac{b}{N}\left(a(t)Q[t,m] - \frac{1}{2}Q^2[t,m]\right) - \frac{1}{2}c_i\left(q_i^S[t,m]\right)^2,$$

$$\pi_i(\mathbf{q}^{S\cup\{i\}}[t,m], \mathbf{q}^{F\setminus\{i\}}[t,m])$$
$$= \frac{b}{N}\left(a(t)Q^{-i}[t,m] - \frac{1}{2}(Q^{-i}[t,m])^2\right) - \frac{1}{2}c_i\left(q_i^{F\cup\{i\}}[t,m]\right)^2,$$

where $Q^{-i}[t,m]$ is the total emission to be reduced over $[t,m]$, if the coalition S is abandoned by one of the players of type i.

Time-consistency of an internally stable agreement is based on the membership preferences of the players. Having accomplished its obligations for the time t, a signatory of type i considers withdrawing from the coalition S, assuming from that moment onwards, the abatement path will continue along the restriction of the optimal solution of the game $\Gamma_0(S\setminus\{i\})$. In such a scenario, a condition for the time-consistency of an internally stable coalition is as follows.

Lemma 6.2.1 *An internally stable agreement in the game $\Gamma_0(S)$ (using Stackelberg equilibrium concept) is time-consistent under the abatement scheme (51), if the following inequality holds for all $t=1,\dots,m-1$, $i=1,\dots,K$,*

$$\Delta_i^S + (2^t-1)\frac{b}{N}(Q-Q^{-i})(a-Q) \geq 0, \tag{52}$$

where [1]

$$\Delta_i^S = \pi_i(\mathbf{q}^S, \mathbf{q}^F) - \pi_i(\mathbf{q}^{S\setminus\{i\}}, \mathbf{q}^{F\cup\{i\}}). \tag{53}$$

P r o o f. Let us analyze a property of time-consistency of internal stability (40) of the agreement S at $t=1,\dots,m-1$, that is

$$\pi_i(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) \geq \pi_i(\mathbf{q}^{S\setminus\{i\}}[t,m], \mathbf{q}^{F\cup\{i\}}[t,m]),$$

for all $i\in S$. The net benefit of a signatory of type i, $i=1,\dots,K$, is

$$\pi_i(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) = \frac{b}{N}\left(a(t)Q[t,m] - \frac{1}{2}Q^2[t,m]\right) - \frac{1}{2}c_i\left(q_i^S[t,m]\right)^2,$$

where

$$a(t) = a - \frac{1}{2}Q - \frac{1}{4}Q - \ldots - \frac{1}{2^t}Q = a - \frac{2^t-1}{2^t}Q.$$

[1] This notation was introduced in Section 4. If the coalition S is self-enforcing, then Δ_i^S is nonnegative.

Then, according to (38),

$$\pi_i(\mathbf{q}^S[t,m],\mathbf{q}^F[t,m]) = \frac{b}{N}\left((a-(1-\frac{1}{2^t})Q)\frac{1}{2^t}Q - \frac{1}{2}\frac{1}{2^{2t}}Q^2\right) - \frac{1}{2}\frac{1}{2^{2t}}c_i\left(q_i^S\right)^2$$
$$= \frac{1}{4^t}\left(\frac{b}{N}(aQ-\frac{1}{2}Q^2) - \frac{1}{2}c_i q_i^{S^2}\right) + \frac{b}{N}(-\frac{1}{2^t}Q^2 + \frac{1}{4^t}Q^2 + \frac{2^t-1}{4^t}aQ)$$
$$= \frac{1}{4^t}\pi_i(\mathbf{q}^S,\mathbf{q}^F) + \frac{2^t-1}{4^t}\frac{b}{N}Q(a-Q).$$

Let us now assume that a signatory of type i leaves the agreement S at a certain moment t. Then its net befit is

$$\pi_i(\mathbf{q}^{S\setminus\{i\}}[t,m],\mathbf{q}^{F\cup\{i\}}[t,m])$$
$$= \frac{b}{N}\left(a(t)Q^{-i}[t,m] - \frac{1}{2}(Q^{-i}[t,m])^2\right) - \frac{1}{2}c_i\left(q_i^{F\cup\{i\}}[t,m]\right)^2,$$

where $Q^{-i}[t,m]$ is restriction over $[t,m]$ of a Stackelberg solution in the game $\Gamma_0(S\setminus\{i\})$. Then

$$\pi_i(\mathbf{q}^S[t,m],\mathbf{q}^F[t,m]) - \pi_i(\mathbf{q}^{S\setminus\{i\}}[t,m],\mathbf{q}^{F\cup\{i\}}[t,m]) \tag{54}$$

$$= \frac{1}{4^t}\left(\pi_i(\mathbf{q}^S,\mathbf{q}^F) - \pi_i(\mathbf{q}^{S\setminus\{i\}},\mathbf{q}^{F\cup\{i\}})\right)$$
$$+ \frac{2^t-1}{4^t}\frac{b}{N}(Q-Q^{-i})(a-Q).$$

Using (53), inequality

$$\pi_i(\mathbf{q}^S[t,m],\mathbf{q}^F[t,m]) \geq \pi_i(\mathbf{q}^{S\setminus\{i\}}[t,m],\mathbf{q}^{F\cup\{i\}}[t,m])$$

can be rewritten as follows

$$\Delta_i^S + (2^t-1)\frac{b}{N}(Q-Q^{-i})(a-Q) \geq 0. \quad \blacksquare$$

Since the condition of internal stability holds at the time $t=0$, then $\Delta_i^S \geq 0$ and the first term in (54) is nonnegative. The sign of the second term in (54)

$$\frac{2^t-1}{4^t}\frac{b}{N}(Q-Q^{-i})(a-Q)$$

is indicated by $Q-Q^{-i}$ (difference between total abatement in cases when agreement is given by set S and $S\setminus\{i\}$) and its sign depends on the model parameters. Let us assume that the agreement is environmentally efficient, *i.e.* inequality (28) holds. The following statement holds.

Theorem 6.2.1 *Assume environmental efficiency condition (28) holds. Then an internally stable agreement in the game $\Gamma_0(S)$ (using the Stackelberg equilibrium concept) is time-consistent under the abatement scheme (51).*

P r o o f. We consider an internally stable agreement S. According to Lemma 6.2.1 condition of time-consistency of internal stability requires

$$\Delta_i^S + \frac{b}{N}(2^t - 1)(Q - Q^{-i})(a - Q) \geq 0.$$

Since the condition of internal stability holds at the time $t = 0$, we see that $\Delta_i^S \geq 0$. The second term

$$(2^t - 1)\frac{b}{N}(Q - Q^{-i})(a - Q)$$

is non-negative as well. Thus during $[0, m]$, an internally stable agreement is time-consistent and no signatory has the incentives to withdraw. ■

Corollary 6.2.1 *Let players of set $\mathcal{N}$ be homogeneous. An internally stable agreement S is time-consistent, if it is composed of at least $\lceil \eta_1 \rceil$ signatories, where η_1 is given in (29)*

$$\eta_1 = \frac{(\lambda + N)(\lambda + 2N + 2\lambda N) - \sqrt{\lambda^2(\lambda + N)^2 + 4N^3(\lambda + N)(\lambda + 1)^2}}{2\lambda(\lambda + N)}.$$

This corollary directly follows from Theorems 3.4.1 and 6.2.1.

Example 6.2.1

The model parameters are assumed to be as in Table 17. We shall examine time-consistency of internal stability of the coalition S_2 (see Table 18) when discounting and pollutant decay rates are $\rho = 1$ and $\theta = 0$, respectively. According to the time-consistent abatement scheme (51), stepwise emission reduction should be as in Table 19.

TABLE 19 Example 6.2.1 Emission Reduction Scheme, m=5.

$[t, t+1)$	$q_1^S = 6.969$	$q_2^S = 4.355$	$q_1^F = 5.11$	$q_2^F = 3.194$
$[0, 1)$	3.485	2.178	2.555	1.597
$[1, 2)$	1.742	1.089	1.278	0.799
$[2, 3)$	0.871	0.544	0.639	0.399
$[3, 4)$	0.436	0.272	0.319	0.2
$[4, 5]$	0.436	0.272	0.319	0.2

Let us confirm that conditions environmental efficiency of the coalition S_2 hold

$$Q^{-1} = 59.615 < Q = 61.672, \quad Q^{-2} = 59.565 < Q = 61.672,$$
$$Q^{+1} = 65.581 > Q = 61.672, \quad Q^{+2} = 64.802 > Q = 61.672.$$

Theorems 6.2.1 and 7.2.1 guarantee that the internal stability of the agreement S is time-consistent upon both free-riding options. This fact is demonstrated in Fig. 9–11.

FIGURE 9 Example 6.2.1. Time-consistency of internal stability of the agreement subject to the players of type 1, $t = 0, \dots, m$, $\rho = 1$, $\theta = 0$. See enraged image in Fig. 10.

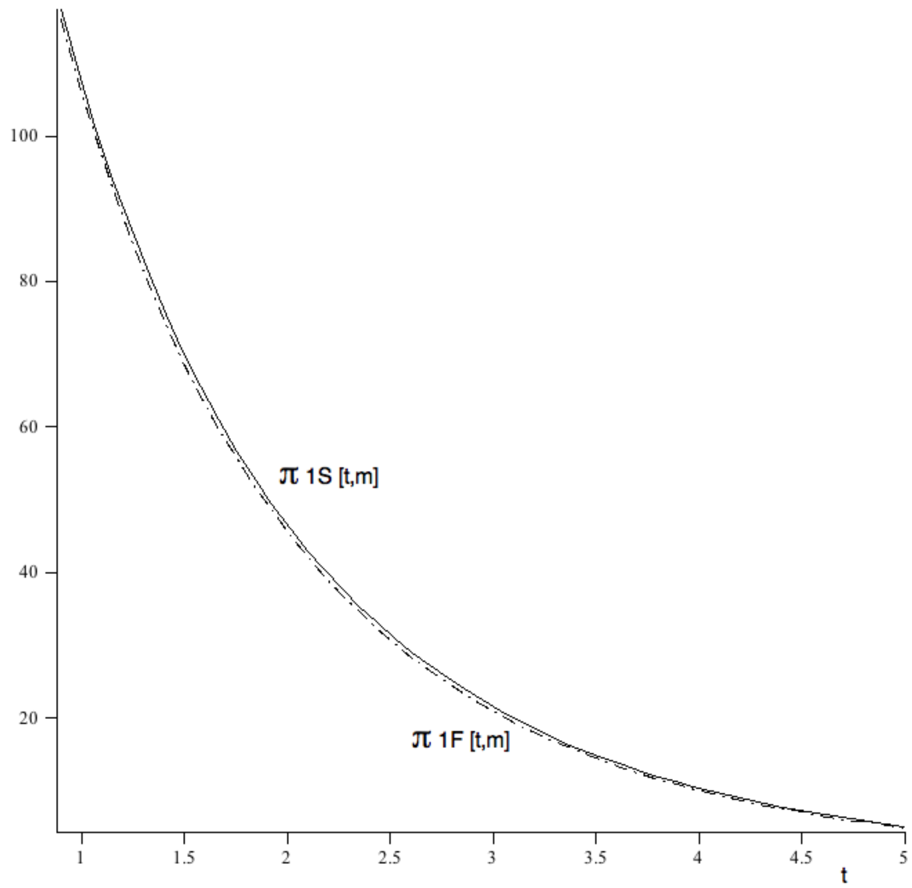


FIGURE 10 Example 6.2.1 . Time-consistency of internal stability of the agreement subject to the players of type 1, $t = 0, \dots, 2$, $\rho = 1$, $\theta = 0$.

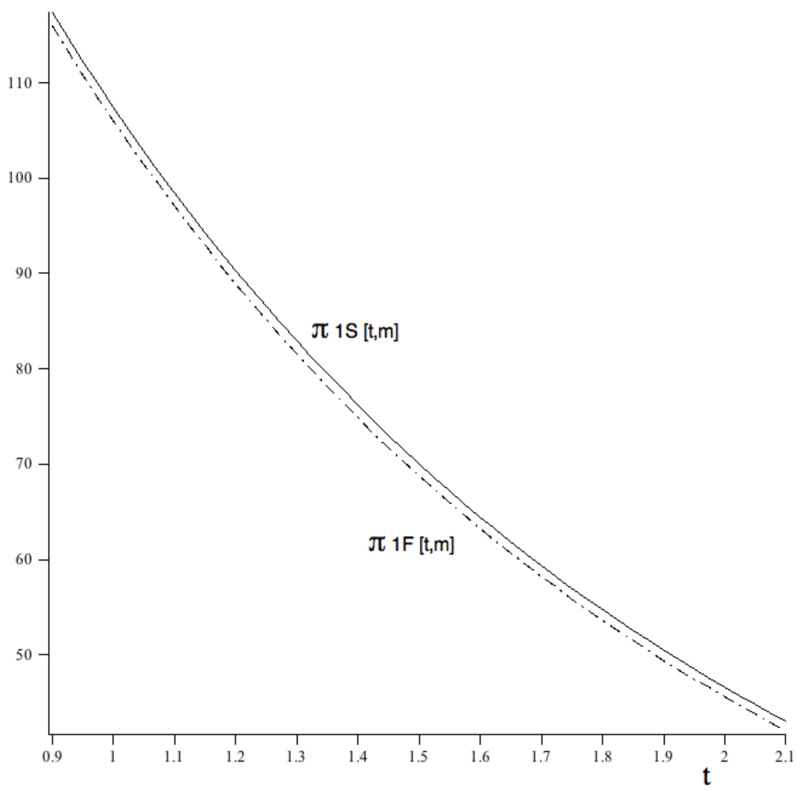
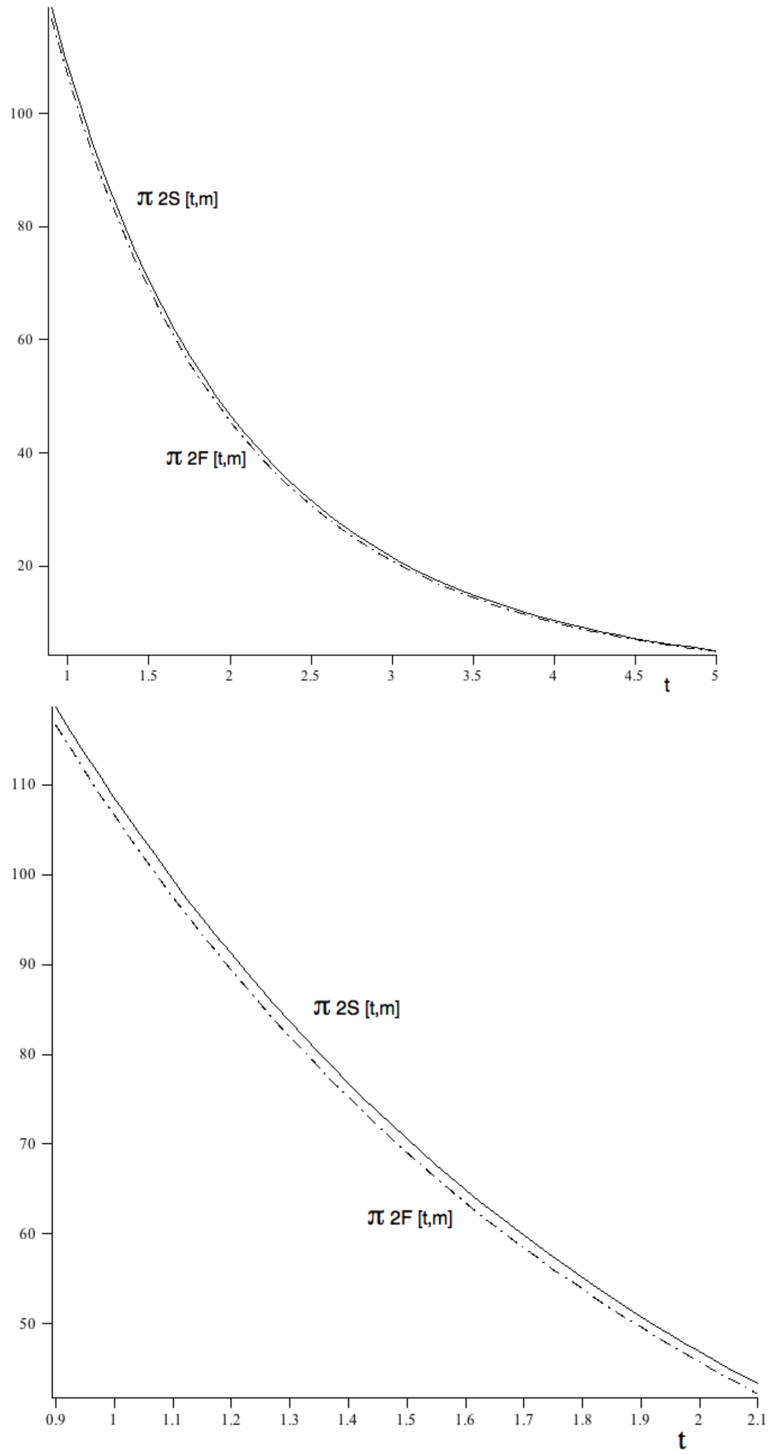


FIGURE 11 Example 6.2.1. Time-consistency of internal stability of the agreement subject to the players of type 2, $t = 0, \dots, m$, $\rho = 1$, $\theta = 0$ (also see enraged image below).



6.3 Time-Consistency of Externally Stable Agreement

Now we are going to analyze the property of time-consistency of external stability of the coalition (Definition 5.0.2) and consider the possibility of joining new nations to the agreement. In Section 6.2 we formulated Lemma 6.2.1 and Theorem 6.2.1, which guaranteed time-consistency of an internally stable agreement. Now we intend to explore external time-consistency and see if the free-riders' attitude towards the agreement can change over time. Given that for time t players reduced emission according to the scheme (51), a free-rider of type i considers joining in the coalition S, assuming that from that moment onwards abatement path will continue along the restriction of the optimal solution of the game $\Gamma_0(S \cup \{i\})$. In such a scenario, we formulate the following statements.

Lemma 6.3.1 *An externally stable agreement in the game $\Gamma_0(S)$ (using Stackelberg equilibrium concept) is time-consistent under the abatement scheme (51) if the following inequality holds for all $t = 1, \ldots, m-1$, $i \in F$,*

$$\Delta_i^{S \cup \{i\}} + (2^t - 1)\frac{b}{N}(a - Q)(Q^{+i} - Q) \leq 0. \tag{55}$$

To verify time-consistency of external stability, we apply a similar approach as earlier in Lemma 6.2.1 and represent the net benefit of a free-rider of type i from the set F as follows

$$\pi_i(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) = \frac{1}{4^t}\pi_i(\mathbf{q}^S, \mathbf{q}^F) + \frac{2^t - 1}{4^t}\frac{b}{N}Q(a - Q).$$

The net benefit of a former free-rider of type i, which accesses the agreement, is

$$\pi_i(\mathbf{q}^{S \cup \{i\}}[t,m], \mathbf{q}^{F \setminus \{i\}}[t,m]) = \frac{1}{4^t}\pi_i(\mathbf{q}^{S \cup \{i\}}, \mathbf{q}^{F \setminus \{i\}}) + \frac{2^t - 1}{4^t}\frac{b}{N}Q^{+i}(a - Q).$$

The condition (41) of time-consistency of external stability is rewritten in the following manner

$$\Delta_i^{S \cup \{i\}} \leq -\frac{b}{N}(2^t - 1)(a - Q)(Q^{+i} - Q).$$

It is useful to remind that

$$\Delta_i^{S \cup \{i\}} = \pi_i(\mathbf{q}^{S \cup \{i\}}, \mathbf{q}^{F \setminus \{i\}}) - \pi_i(\mathbf{q}^S, \mathbf{q}^F), \qquad \forall i \in F,$$

due to the assumption of external stability of the coalition S in the game $\Gamma_0(S)$

Theorem 6.3.1 *Assume the condition (28) of environmental efficiency holds. The abatement scheme (51) does not guarantee time-consistency of an externally stable agreement S in the game $\Gamma_0(S)$.*

P r o o f. To verify the time-consistency condition of external stability, we are going to use inequality (55)

$$\Delta_i^{S\cup\{i\}} + \frac{b}{N}(2^t - 1)(a - Q)(Q^{+i} - Q) \leq 0, \qquad i \in F.$$

The condition (28) of environmental efficiency holds, then (55) can be represented as follows

$$\begin{aligned} & -\Delta_i^{S\cup\{i\}} + \frac{b}{N}(a - Q)(Q^{+i} - Q) \geq 2^t \frac{b}{N}(a - Q)(Q^{+i} - Q), \\ & 2^t \leq \frac{-\Delta_i^{S\cup\{i\}} + \frac{b}{N}(a - Q)(Q^{+i} - Q)}{\frac{b}{N}(a - Q)(Q^{+i} - Q)}, \\ & t \leq \log_2 \frac{-\frac{N}{b}\Delta_i^{S\cup\{i\}} + (a - Q)(Q^{+i} - Q)}{(a - Q)(Q^{+i} - Q)}. \end{aligned}$$

Herewith we conclude that external time-consistency holds for $t < t^*$, where

$$t^* = \left[\log_2 \frac{-\frac{N}{b}\Delta_i^{S\cup\{i\}} + (a - Q)(Q^{+i} - Q)}{(a - Q)(Q^{+i} - Q)}\right] + 1 \tag{56}$$

determines a threshold stage. Notation $[\ldots]$ means an integer part of a number. Notice that since

$$-\frac{N}{b}\Delta_i^{S\cup\{i\}} + (a - Q)(Q^{+i} - Q) \geq (a - Q)(Q^{+i} - Q),$$

then

$$\frac{-\frac{N}{b}\Delta_i^{S\cup\{i\}} + (a - Q)(Q^{+i} - Q)}{(a - Q)(Q^{+i} - Q)} > 1,$$

and hence $t^* \geq 1$. As soon as $t \geq t^*$, external time-consistency is violated. If t^* is less than two then external time-consistency breaks at the stage $t = 1$, if it is less than three then external time-consistency breaks at the stage $t = 2$ and so on. ■

Example 6.3.1

Let us reconstruct the example, given in [Barrett 1994a], considering coalition formation process among 10 homogeneous players (see also Example 3.4.1). We remind that according to [Barrett 1994a] parameters of the game $\Gamma_0(S)$ are as follows: $N = 10$, $a = 100$, $b = 1$, and $c = 0.25$. The process of identifying a self-enforcing coalition delivers a single stable IEA composed of four signatories (see Table 20).

This agreement, if formed, would provide total emission reduction $Q = 81.069$ (of $a = 100$), payoff of signatories and free-riders are $\pi_{i\in S} = 472.16$ and $\pi_{i\in F} = 474.913$, respectively. Environmental target will be met in line with the following abatement scheme (51)

$$q^S[t, m] = \frac{1}{2^t}8.909, \qquad q^F[t, m] = \frac{1}{2^t}7.572,$$

TABLE 20 Example 6.3.1 Coaliton Stability Analysis.

n	q_i^S	q_i^F	π_S	π_F	Q	Π
0	0	8	0	472	80	4720
1	1.855	8.534	476.809	468.135	78.664	4690.022
2	4.158	8.732	474.012	466.643	78.170	4681.169
3	6.652	8.426	472.284	468.941	78.936	4699.436
4*	8.909	7.572	472.160	474.913	81.069	4738.121
5	10.526	6.316	473.684	482.548	84.211	4781.163
6	11.342	4.915	476.371	489.431	87.713	4815.949
7	11.457	3.6	479.542	494.328	90.998	4839.776
8	11.096	2.497	482.663	497.273	93.759	4855.85
9	10.477	1.63	485.448	498.838	95.925	4867.872
10	9.756	0	487.805	0	97.56	4878.049

where $t = 0, m-1$. Let $m = 10$ then the agreement commitments $(\mathbf{q}^S, \mathbf{q}^F) = (\mathbf{8.909}, \mathbf{7.572})$ should be fulfilled over 10 steps. Following the time-consistent abatement scheme, players perform emission reduction during each time period according to Table 21.

TABLE 21 Example 6.3.1 Players"Per-Interval" Abatement Plans.

Intervals $\rightarrow$ *Example 6.3.1. Players* $\downarrow$	$[0,1)$	$[1,2)$	$[2,3)$	$[3,4)$	$[4,5)$	$[5,6)$	$[6,7)$
'signatory'	4.455	2.227	1.114	0.557	0.278	0.139	0.07
'free-rider'	3.786	1.893	0.947	0.473	0.237	0.118	0.059

Intervals $\rightarrow$ *Example 6.3.1. Players* $\downarrow$	$[7,8)$	$[8,9)$	$[9,10]$
'signatory'	0.035	0.017	0.017
'free-rider'	0.03	0.015	0.015

Let us calculate $\eta_{1,2}$, determined in (29),

$$\eta_1 = 2.424, \qquad \eta_2 = 23.576.$$

According to Theorem 3.4.1 we can guarantee that the agreement is environmentally efficient when it is composed of $\lceil \eta_1 \rceil = 3$ and more players. From Table 20 we conclude that for $n = 0,1,2$ $Q(n) - Q(n-1) < 0$ and for $n = 3,\dots,N$ $Q(n) - Q(n-1) > 0$. Consequently, a coalition which consists of four satisfies condition of environmental efficiency, and, according to Lemma 6.2.1 and Theorem 6.2.1, the internally stable agreement is time-consistent over $t = 0,\dots,m$. It means that no agreement members would unilaterally abandon the coalition during $[0, m]$. Using the abatement scheme (51) guarantees time-consistency of internal stability of the coalition S (Theorem 6.2.1) but fails to sustain time-consistency of external stability of the coalition after a threshold moment t^*, determined according to (56) and equal to 1 in the current example.

6.4 Total Payoffs over Accounting Periods

We assume that the net present value of each abating interval is decreasing, as it occurs due to natural uncertainty about the future. In practice, it means that profits to be received in a certain time interval, are less valuable in comparison to those, we can get 'today'.

Usage of the net present value notation requires the discount factor

$$\rho = \left(1 + \frac{r}{m}\right)^{-1},$$

where r is an interest rate. Traditionally $r \ll 1$, for instance, $r = 0.05$. Then it follows that $0 < \rho < 1$.

Previously we have constructed the time-consistent abatement scheme and considered property of time-consistency of a self-enforcing coalition. Now we are going to see what happens to the players' total net benefit in dynamics using such a scheme

$$\sum_{t=0}^{m-1} \rho^t \pi_i(\Delta\mathbf{q}^S[t, t+1), \Delta\mathbf{q}^F[t, t+1)). \tag{57}$$

Unfortunately, it is impossible to evaluate analytically (57) for the scheme (51) due to high complexity of the expressions.

Let us consider a particular case of scheme (51), where $\rho = 1$ and $\theta = 0$. We are going to summarize received payoffs from the moment $t = 1$ to $t = m$ and compare results to the net benefits assigned to players at the moment $t = 0$, when formation of the agreement structure and abatement goals takes place. We are also going stepwise to increase the number of check-points (and thus increase the number of periods $[0, 1), \ldots, [m-1, m]$) to observe changes in players' payoffs and provide assessment of the length of the abating period.

- If $m = 1$ then individual payoff of a player of type i, according to (1), is

  $$\pi_i(\mathbf{q}^S, \mathbf{q}^F) = \frac{b}{N}(aQ - \frac{1}{2}Q^2) - \frac{1}{2}c_i \left(q_i^{S(F)}\right)^2,$$

  this is also the anticipated payoff for the player.

- If $m = 2$ then individual payoffs obtained by players over the intervals $[0, 1)$ and $[1, 2]$, are

  $$\sum_{t=0}^{1} \pi_i(\mathbf{q}^S[t, t+1), \mathbf{q}^F[t, t+1)) = \frac{b}{N}(a\frac{1}{2}Q - \frac{1}{2}(\frac{1}{2}Q)^2) - \frac{1}{2}c_i \left(\frac{1}{2}q_i^{S(F)}\right)^2$$

$$+\frac{b}{N}((a-\frac{1}{2}Q)\frac{1}{2}Q-\frac{1}{2}(\frac{1}{2}Q)^2)-\frac{1}{2}c_i\left(\frac{1}{2}q_i^{S(F)}\right)^2$$
$$=\frac{b}{N}(aQ-\frac{1}{2}Q^2)-\frac{1}{4}c_i\left(q_i^{S(F)}\right)^2=\pi_i^{S(F)}+\frac{1}{4}c_i\left(q_i^{S(F)}\right)^2.$$

As one may see, the finally received payoffs of signatories and free-riders are bigger than those anticipated during agreement formation. This occurs due to model structure and split of the total abatement into two parts. At the same time it serves as an incentive to follow the dynamic scheme.

- If $m=3$ then individual payoffs over $[0,1)$,$[1,2)$ and $[2,3]$

$$\sum_{t=0}^{2}\pi_i(\mathbf{q}^S[t,t+1),\mathbf{q}^F[t,t+1))=\frac{b}{N}(a\frac{1}{2}Q-\frac{1}{2}(\frac{1}{2}Q)^2)-\frac{1}{2}c_i\left(\frac{1}{2}q_i^{S(F)}\right)^2$$
$$+\frac{b}{N}((a-\frac{1}{2}Q)\frac{1}{4}Q-\frac{1}{2}(\frac{1}{4}Q)^2)-\frac{1}{2}c_i\left(\frac{1}{4}q_i^{S(F)}\right)^2$$
$$+\frac{b}{N}((a-\frac{1}{2}Q-\frac{1}{4}Q)\frac{1}{4}Q-\frac{1}{2}(\frac{1}{4}Q)^2)-\frac{1}{2}c_i\left(\frac{1}{4}q_i^{S(F)}\right)^2$$
$$=\frac{b}{N}(aQ-\frac{1}{2}Q^2)-\frac{3}{16}c_i\left(q_i^{S(F)}\right)^2=\pi_i(\mathbf{q}^S,\mathbf{q}^F)+\frac{5}{16}c_i\left(q_i^{S(F)}\right)^2.$$

In this case, the final received payoffs of signatories and free-riders are also bigger than those anticipated during agreement formation, which is a stronger incentive for players to accept this scheme. At the same time one may notice that additional profit for $m=3$ is higher than the one for $m=1$, which is $\frac{5}{16}\geq\frac{1}{4}$, and

$$\sum_{t=0}^{2}\pi_i(\mathbf{q}^S[t,t+1),\mathbf{q}^F[t,t+1))-\sum_{t=0}^{1}\pi_i(\mathbf{q}^S[t,t+1),\mathbf{q}^F[t,t+1))=\frac{1}{16}>0.$$

- If $m=4$ then individual payoffs over $[0,1)$, $[1,2)$, $[2,3)$ and $[3,4]$ are

$$\sum_{t=0}^{3}\pi_i(\mathbf{q}^S[t,t+1),\mathbf{q}^F[t,t+1))$$
$$=\frac{b}{N}\left(a\frac{1}{2}Q-\frac{1}{2}(\frac{1}{2}Q)^2\right)-\frac{1}{2}c_i\left(\frac{1}{2}q_i^{S(F)}\right)^2$$
$$+\frac{b}{N}\left((a-\frac{1}{2}Q)\frac{1}{4}Q-\frac{1}{2}(\frac{1}{4}Q)^2\right)-\frac{1}{2}c_i\left(\frac{1}{4}q_i^{S(F)}\right)^2$$
$$+\frac{b}{N}\left((a-\frac{1}{2}Q-\frac{1}{4}Q)\frac{1}{8}Q-\frac{1}{2}(\frac{1}{8}Q)^2\right)-\frac{1}{2}c_i\left(\frac{1}{8}q_i^{S(F)}\right)^2$$
$$+\frac{b}{N}\left((a-\frac{1}{2}Q-\frac{1}{4}Q-\frac{1}{8}Q)\frac{1}{8}Q-\frac{1}{2}(\frac{1}{8}Q)^2\right)-\frac{1}{2}c_i\left(\frac{1}{8}q_i^{S(F)}\right)^2$$
$$=\frac{b}{N}\left(aQ-\frac{1}{2}Q^2\right)-\frac{11}{64}c_i\left(q_i^{S(F)}\right)^2=\pi_i\left(\mathbf{q}^S,\mathbf{q}^F\right)+\frac{21}{64}c_i\left(q_i^{S(F)}\right)^2.$$

When $m = 4$ total net benefits are higher than it was anticipated at the initial moment and besides that

$$\sum_{t=0}^{3} \pi_i(\mathbf{q}^S[t,t+1),\mathbf{q}^F[t,t+1)) - \sum_{t=0}^{2} \pi_i(\mathbf{q}^S[t,t+1),\mathbf{q}^F[t,t+1)) = \frac{1}{64} > 0.$$

- Individual payoffs over $[0,1),[1,2),\dots,[m-1,m]$ are

$$\sum_{t=0}^{m-1} \pi_i(\mathbf{q}^S[t,t+1),\mathbf{q}^F[t,t+1))$$

$$= \sum_{t=1}^{m-1}\left[\frac{b}{N}\left((a-\sum_{j=1}^{t-1}\frac{1}{2^j}Q)\frac{1}{2^t}Q - \frac{1}{2}\frac{1}{4^t}Q^2\right) - \frac{1}{2}c_i\frac{1}{4^t}(q_i^{S(F)})^2\right] \quad (58)$$

$$+\frac{b}{N}\left((a-\sum_{j=1}^{m-1}\frac{1}{2^j}Q)\frac{1}{2^{m-1}}Q - \frac{1}{2}\frac{1}{4^{m-1}}Q^2\right) - \frac{1}{2}c_i\frac{1}{4^{m-1}}\left(q_i^{S(F)}\right)^2.$$

In order to obtain general form for individual total payoffs of the players and assess properties of function $\pi_i^{S(F)}(m)$, we bring the expression (58) to the simplified presentation

$$\sum_{t=0}^{m-1} \pi_i(\mathbf{q}^S[t,t+1),\mathbf{q}^F[t,t+1))$$

$$= \underbrace{\frac{b}{N}aQ\sum_{t=1}^{m-1}\frac{1}{2^t}}_{\text{1) term}} - \underbrace{\frac{b}{N}Q^2\sum_{t=1}^{m-1}\frac{1}{2^t}\sum_{j=1}^{t-1}\frac{1}{2^j}}_{\text{2) term}} - \underbrace{\frac{b}{N}\frac{1}{2}Q^2\sum_{t=1}^{m-1}\frac{1}{4^t}}_{\text{3) term}}$$

$$+\underbrace{\frac{b}{N}aQ\frac{1}{2^{m-1}}}_{\text{4) term}} - \underbrace{\frac{b}{N}\frac{1}{2^{m-1}}Q^2\sum_{t=1}^{m-1}\frac{1}{2^t}}_{\text{5) term}} - \underbrace{\frac{1}{2}\frac{1}{4^{m-1}}Q^2}_{\text{6) term}}$$

$$-\underbrace{\frac{1}{2}c_i\left(q_i^{S(F)}\right)^2\sum_{t=1}^{m-1}\frac{1}{4^t}}_{\text{7) term}} - \underbrace{\frac{1}{2}\frac{1}{4^{m-1}}c_iq_i^{S(F)}}_{\text{8) term}}.$$

Consider the first term

$$1): \qquad \frac{b}{N}aQ\sum_{t=1}^{m-1}\frac{1}{2^t} = \frac{b}{N}aQ\frac{\frac{1}{2}(1-\frac{1}{2^{m-1}})}{\frac{1}{2}} = \frac{b}{N}aQ(1-\frac{1}{2^{m-1}}).$$

Now sum it up with the forth term

$$1)+4): \qquad \frac{b}{N}aQ(1-\frac{1}{2^{m-1}}) + \frac{b}{N}aQ\frac{1}{2^{m-1}} = \frac{b}{N}aQ.$$

Consider the second, the third and the fifth terms consequently

$$2): \qquad \frac{b}{N}Q^2 \sum_{t=1}^{m-1} \frac{1}{2^t} \sum_{j=1}^{t-1} \frac{1}{2^j} = \frac{b}{N}Q^2 \sum_{t=1}^{m-1} \frac{1}{2^t}(1 - \frac{1}{2^{t-1}})$$

$$= \frac{b}{N}Q^2 \sum_{t=1}^{m-1} \frac{1}{2^t} - \frac{b}{N}Q^2 \sum_{t=1}^{m-1} \frac{2}{4^t} = \frac{b}{N}Q^2 \left((1 - \frac{1}{2^{m-1}}) - (\frac{2}{3} - \frac{2}{3}\frac{1}{4^{m-1}}) \right)$$

$$\frac{b}{N}Q^2 \left(\frac{1}{3} - \frac{1}{2^{m-1}} + \frac{2}{3}\frac{1}{4^{m-1}} \right),$$

$$3): \qquad \frac{b}{N}\frac{1}{2}Q^2 \sum_{t=1}^{m-1} \frac{1}{4^t} = \frac{b}{N}\frac{1}{2}Q^2 \frac{1}{3}\left(1 - \frac{1}{4^{m-1}}\right) = \frac{b}{N}Q^2 \left(\frac{1}{6} - \frac{1}{6}\frac{1}{4^{m-1}} \right),$$

$$5): \qquad \frac{b}{N}\frac{1}{2^{m-1}}Q^2 \sum_{t=1}^{m-1} \frac{1}{2^t} = \frac{b}{N}\frac{1}{2^{m-1}}Q^2 \left(1 - \frac{1}{2^{m-1}}\right)$$

$$= \frac{b}{N}Q^2 \left(\frac{1}{2^{m-1}} - \frac{1}{4^{m-1}} \right).$$

Now we sum simplified expressions $2), 3), 5)$ and $6)$

$$2) \quad + \quad 3) \quad + \quad 5) + \quad 6):$$

$$\frac{b}{N}Q^2 \left(\frac{1}{3} - \frac{1}{2^{m-1}} + \frac{2}{3}\frac{1}{4^{m-1}} \right) + \frac{b}{N}Q^2 \left(\frac{1}{6} - \frac{1}{6}\frac{1}{4^{m-1}} \right)$$

$$+\frac{b}{N}Q^2 \left(\frac{1}{2^{m-1}} - \frac{1}{4^{m-1}} \right) + \frac{1}{2}\frac{1}{4^{m-1}}Q^2$$

$$= \frac{b}{N}\frac{Q^2}{2}.$$

Consider the last two terms

$$\frac{1}{2}c_i \left(q_i^{S(F)}\right)^2 \sum_{t=1}^{m-1} \frac{1}{4^t} + \frac{1}{2}\frac{1}{4^{m-1}}c_i q_i^{S(F)}$$

$$= \frac{1}{2}c_i \left(q_i^{S(F)}\right)^2 \left((\frac{1}{3} - \frac{1}{3}\frac{1}{4^{m-1}}) + \frac{1}{4^{m-1}} \right)$$

$$= \frac{1}{2}c_i \left(q_i^{S(F)}\right)^2 \left(\frac{1}{3} + \frac{2}{3}\frac{1}{4^{m-1}} \right).$$

Bringing all terms together we obtain that

$$\sum_{t=0}^{m-1} \pi_i(\mathbf{q}^S[t, t+1), \mathbf{q}^F[t, t+1))$$

$$= \frac{b}{N}\left(aQ - \frac{1}{2}Q^2\right) - \frac{1}{2}c_i \left(q_i^{S(F)}\right)^2 \left(\frac{1}{3} + \frac{2}{3}\frac{1}{4^{m-1}} \right) \tag{59}$$

$$= \sum_{t=0}^{m-1} \pi_i(\mathbf{q}^S, \mathbf{q}^F) + \frac{1}{3} c_i \left(q_i^{S(F)} \right)^2 \left(1 - \frac{1}{4^{m-1}} \right).$$

When m is approaching infinity, players' payoffs

$$\lim_{m \to \infty} \sum_{t=0}^{m-1} \pi_i(\mathbf{q}^S[t, t+1], \mathbf{q}^F[t, t+1])$$

$$= \pi_i(\mathbf{q}^S, \mathbf{q}^F) + \frac{1}{3} c_i \left(q_i^{S(F)} \right)^2 \left(1 - \lim_{m \to \infty} \frac{1}{4^{m-1}} \right) = \pi_i(\mathbf{q}^S, \mathbf{q}^F) + \frac{1}{3} c_i \left(q_i^{S(F)} \right)^2.$$

This occurrence of the additional term $\frac{1}{3} c_i \left(q_i^{S(F)} \right)^2$ can be explained by specification of our model (quadratic structure of cost function) and the well-known Cauchy-Schwarz inequality. Such positive surplus $\frac{1}{3} c_i \left(q_i^{S(F)} \right)^2$ can be treated as an additional incentive towards multi-stage abatement process in comparison to static case. On the other hand, assuming that players are uncertain about their future and possible forthcoming events, this additional surplus will obviously become extinct with total individual payoffs being brought to present value.

Material of the current section has been a consequent conclusion of Part II, which main objective was to focus on dynamic framework of the environmental agreement and in particular, to cover two crucial topics: reallocation of emission reduction commitments over a number of time periods; and, analysis of agreement stability upon pollution dynamics. We have suggested a time-consistent abatement reallocation scheme that corresponds to the optimality concept of Stackelberg equilibrium, applied in the game $\Gamma_0(S)$. It assigns emission reductions so that choice of abatement efforts during each time period is adjusted according to the emission reduction undertaken previously, plus taking into account the current environmental settings.

While analyzing time-consistency of a self-enforcing agreement, we have mainly concentrated our attention on time-conistency of internal stability, assuming that after formation of an agreement further accession of new members is prohibited. We have shown that internal stability is time-consistent upon the time-consistent abatement scheme and no signatory has incentives to leave the agreement as collaboration develops. This result can successfully be applied for the case when during the IEA formation all the players are attracted into the agreement (the grand coalition) or the most relevant players access the coalition.

Here we came to a conclusion that, in dynamic framework, external stability never holds over the whole accounted period, which means that after a threshold stage the coalition can be accessed by new members. This signals about both positive and negative perspectives. On the one hand a larger coalition can result in bigger emission reduction and thus bigger environmental benefit. On the other hand time-inconsistency of the agreement means vulnerability of agreement stability in dynamic framework and may also lead to the opposite results, such as full or partial decomposition of the agreement and low environmental benefit. In the following part we are going to continue time-consistency analysis of the self-enforcing coalition and suggest renegotiation mechanisms, which targets to enliven the agreements, whose stability is in jeopardy.

Part III

Renegotiation

7 MECHANISMS OF DYNAMIC REGULARIZATION

7.1 Renegotiation of Self-Enforcing Agreement

As Lemma 6.2.1 and Theorem 6.2.1 point out, we can guarantee that internal time-consistency of a self-enforcing coalition holds under assumptions of environmental efficiency (28). Lemma 6.3.1 and Theorem 6.3.1 state that external time-consistency is violated after a certain threshold stage t^*, which means that at the stage t^* the agreement requires renegotiation. Material, presented in Part III, is based on publications [Pavlova 2008], [Pavlova *et al.* 2008]. Here we consider the renegotiation process, which occurs in the dynamic game $\Gamma_{t^*}\left(S, \mathbf{q}^S[0,t^*), \mathbf{q}^F[0,t^*)\right)$ and can be interpreted as repetition of a one-shot game $\Gamma_0(S)$ but with different environmental setting $a(t^*)$.

Let us further suppose that emission reduction process goes along with the time-consistent abatement scheme (51). At t^* players simultaneously and voluntary choose their status towards the IEA (signatory/free-rider). Agreement signatories $\mathbf{n}^* = \left(n_1^*, \ldots, n_i^*, \ldots, n_K^*\right)$ make decisions about their emission reduction commitments $q_i^{S_{t^*}}$, which must be fulfilled during $[t^*, m]$, by maximizing aggregate coalition net benefit $\sum_{i \in S_{t^*}} \pi_i(\mathbf{q}^{S_{t^*}}, \mathbf{q}^{F_{t^*}})$. Free-riders from set $F_{t^*} = \mathcal{N} \setminus S_{t^*}$ adjust their abatement levels, taking the choice of signatories as given, and each of them maximizes its net benefit $\pi_i(\mathbf{q}^{S_{t^*}}, \mathbf{q}^{F_{t^*}})$, $i \in F$, non-cooperatively.

As before we suppose parameters c_i and b constant. Current environmental situation is characterized by parameter $a(t^*)$

$$a(t^*) = a(t^*-1) - \frac{1}{2^{t^*}}Q = a - \frac{2^{t^*}-1}{2^{t^*}}Q;$$

this formula follows from (38) and (51). In order to identify structure of the self-enforcing coalition S_{t^*}, it is necessary to redefine abatement strategies $\left(\mathbf{q}^{S_{t^*}}, \mathbf{q}^{F_{t^*}}\right)$.

In the two level game $\Gamma_{t^*}(S_{t^*})$ Stackelberg equilibrium is unique and constituted by the following strategies of the leader (the coalition S_{t^*})

$$q_i^{S_{t^*}} = \frac{a(t^*)\lambda_i(1-g^*)^2(\bar{1}, \mathbf{n}^*)}{(\bar{1}, \mathbf{N}) + (1-g^*)^2(\bar{1}, \mathbf{n}^*)(\lambda, \mathbf{n}^*)}, \quad i = 1, \ldots, K, \tag{60}$$

and the followers (the free-riders from set F_{t^*}),

$$q_i^{F_{t^*}} = \frac{\lambda_i a(t^*)(\bar{1}, \mathbf{N})}{[(\bar{1}, \mathbf{N}) + (1-g)^2(\bar{1}, \mathbf{n}^*)(\lambda, \mathbf{n}^*)][(\bar{1}+\lambda, \mathbf{N}) - (\lambda, \mathbf{n}^*)]}, \quad i = 1, \ldots, K, \tag{61}$$

where

$$g^* = \frac{(\lambda, \mathbf{N} - \mathbf{n}^*)}{N + (\lambda, \mathbf{N} - \mathbf{n}^*)}.$$

Functions $\left(\mathbf{q}^{S_{t^*}}, \mathbf{q}^{F_{t^*}}\right)$ should be placed into conditions of a self-enforcing coalition, which can be rewritten as follows

$$\pi_i(\mathbf{q}^{S_{t^*}}, \mathbf{q}^{F_{t^*}}) \geq \pi_i(\mathbf{q}^{S_{t^*} \setminus \{i\}}, \mathbf{q}^{F_{t^*} \cup \{i\}}), \quad i \in S_{t^*},$$

$$\pi_i(\mathbf{q}^{S_{t^*} \cup \{i\}}, \mathbf{q}^{F_{t^*} \setminus \{i\}}) \leq \pi_i(\mathbf{q}^{S_{t^*}}, \mathbf{q}^{F_{t^*}}), \quad i \in F_{t^*}.$$

Solution of this system under the found expressions (60), (61) describes structure of the self-enforcing agreement S_{t^*}. Renegotiation process possesses the following properties:

- Set of self-enforcing coalitions, determined upon Stackelberg solution $(\mathbf{q}^{S_{t^*}}, \mathbf{q}^{F_{t^*}})$ (60), (61) during renegotiation process at the stage t^* in the game $\Gamma_{t^*}(S_{t^*})$, is equivalent to set, obtained upon Stackelberg solution $(\mathbf{q}^S, \mathbf{q}^F)$ (19), (20) of the game $\Gamma_0(S)$.

  Identity of two sets of self-enforcing coalitions follows from the similar structure of abatement solutions $\left(\mathbf{q}^{S_{t^*}}, \mathbf{q}^{F_{t^*}}\right)$ at stage t^* and $\left(\mathbf{q}^S, \mathbf{q}^F\right)$ at stage 0, where the main difference is in parameter $a(t)$, which characterizes environmental situation. In case there are two or more self-enforcing coalitions possible, it is possible that the appeared IEA has the same structure $\mathbf{n} = (n_1, \ldots, n_i, \ldots, n_K)$ of the coalition as it was at the initial stage $t = 0$.

  Redefined coalition continues abatement process. We suppose that the emission reduction dynamics is held according to the time-consistent abatement scheme (51), suggested in Section 6.1 and described as a geometric progression with 0.5 as a common ratio and $\left(\mathbf{q}^{S_{t^*}}, \mathbf{q}^{F_{t^*}}\right)$ as an initial element, *i.e.* for $t^* \leq t \leq m-1$

  $$\Delta q_i^{S_{t^*}}[t, t+1) = \frac{1}{2^{t-t^*+1}} q_i^{S_{t^*}},$$

  $$\Delta q_i^{F_{t^*}}[t, t+1) = \frac{1}{2^{t-t^*+1}} q_i^{F_{t^*}}.$$

- In the multistage game, renegotiation is a regular process, occurring with a period t^* starting from the initial moment 0.

  This statement can be straightforwardly obtained by placing $t = t^*$ as the initial moment of time-consistency analysis, and obtaining that the new threshold level is equal to t^*. It implies that the following renegotiations will take place at stages $2t^*$, $3t^*$, $4t^*$ and so on until $m-1$, and emission reduction during time period $[jt^*, (j+1)t^*]$ will be held as follows

$$\Delta q_i^{S_{t^*}}[t, t+1) = \frac{1}{2^{t-t^*+1}} q_i^{S_{t^*}}, \tag{62}$$

$$\Delta q_i^{F_{t^*}}[t, t+1) = \frac{1}{2^{t-t^*+1}} q_i^{F_{t^*}},$$

when $jt^* \le t \le (j+1)t^* - 1$. Here $j = 1, 2, 3, \ldots$, so that $(j+1)t^* - 1 \le m$. Strategies $(\mathbf{q}^{S_{t^*}}, \mathbf{q}^{F_{t^*}})$ are given in (60), (61).

- Specifying abatement dynamics according to (62), a self-enforcing coalition S, which is self-enforcing in the game $\Gamma_0(S)$ upon Stackelberg equilibrium concept, is time-consistent.

Example 7.1.1

Let us consider the example, given in [Barrett 1994a], describing the process of coalition formation among homogeneous players, which results in a single self-enforcing IEA. According to [Barrett 1994a] parameters of the game $\Gamma_0(S)$ are as follows: $N = 10$, $a = 100$, $b = 1$, and $c = 0.25$. The process of identifying a self-enforcing coalition, described in Section 3.3.3, delivers a single stable IEA composed of four signatories (see Table 22).

TABLE 22 Example 7.1.1 Self-Enforcing Coalitions.

n	q_i^S	q_i^F	π^S	π^F	Q	Π
0	0	8	0	472	80	4720
1	1.855	8.534	476.809	468.135	78.664	4690.022
2	4.158	8.732	474.012	466.643	78.170	4681.169
3	6.652	8.426	472.284	468.941	78.936	4699.436
4*	8.909	7.572	472.160	474.913	81.069	4738.121
5	10.526	6.316	473.684	482.548	84.211	4781.163
6	11.342	4.915	476.371	489.431	87.713	4815.949
7	11.457	3.6	479.542	494.328	90.998	4839.776
8	11.096	2.497	482.663	497.273	93.759	4855.85
9	10.477	1.63	485.448	498.838	95.925	4867.872
10	9.756	0	487.805	0	97.56	4878.049

Let the agreement commitments $(\mathbf{q}^S, \mathbf{q}^F) = (\mathbf{8.909}, \mathbf{7.572})$ be fulfilled over five steps. Following the time-consistent abatement scheme, players perform emission reduction during each time period according to Table 23.

The coalition S, composed of four players, satisfies property of environmental efficiency (28) since aggregate emission reduction $Q = 81.069$ is larger than total emission reduction $Q^{-i} = 78.936$ for the coalition, composed of only three

TABLE 23 Example 7.1.1 Emission Reduction Scheme.

Time period $\rightarrow$ *Individual abatements*$\downarrow$	$[0,1)$	$[1,2)$	$[2,3)$	$[3,4)$	$[4,5]$
$\Delta q_i^S[t,t+1)$	4.455	2.227	1.114	0.557	0.557
$\Delta q_i^F[t,t+1)$	3.786	1.893	0.947	0.473	0.473

players. Using the abatement scheme (51) guarantees time-consistency of internal stability of the coalition S (Theorem 6.2.1) but fails to sustain time-consistency of external stability of the coalition after a threshold moment t^*, determined according to (56) and equal to 1 in the current example. It implies that the renegotiation process, described in Section 7.1, will occur after the first abatement period. Model parameters at $t = 1$ are described in Table 24. Pollution flow evolves according to (38) and $a(1) = a - \frac{1}{2}E = 60.4655$.

TABLE 24 Example 7.1.1 Model Parameters at the First Renegotiation.

$b = 1$	$c = 0.25$
$a(1) = 60.4655$	$N = 10$

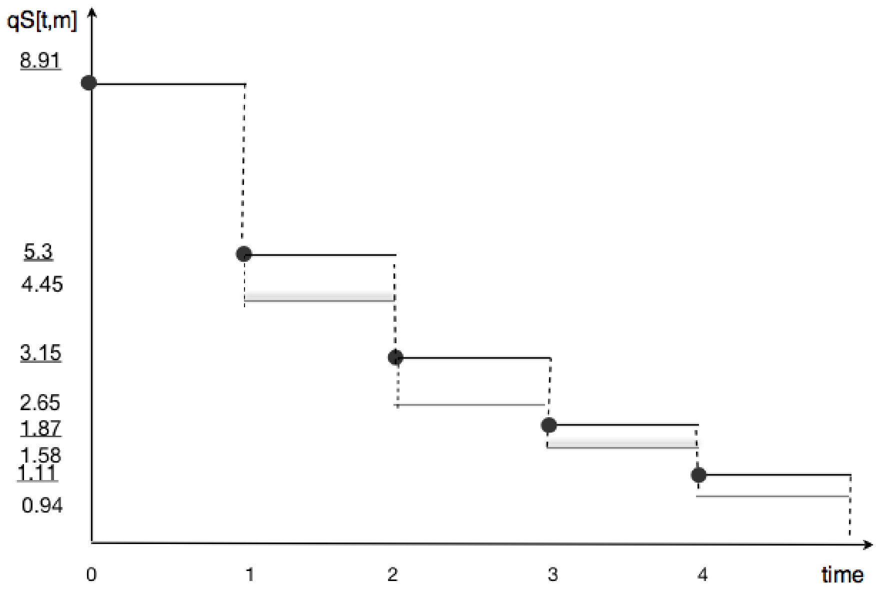
TABLE 25 Example 7.1.1 Self-Enforcing Coalition after the First Renegotiation.

n	q_i^S	q_i^F	π^S	π^F	Q	Π
0	0	4.837	0	172.567	48.372	1725.668
1	1.122	5.160	174.325	171.154	47.565	1714.708
2	2.514	5.28	173.303	170.608	47.266	1711.471
3	4.022	5.095	172.671	171.448	47.729	1718.15
4*	5.387	4.579	172.625	173.632	49.019	1732.293
5	6.365	3.819	173.183	176.423	50.918	1748.03
6	6.858	2.972	174.165	178.94	53.036	1760.748
7	6.927	2.177	175.324	180.73	55.023	1769.459
8	6.709	1.51	176.465	181.807	56.692	1775.336
9	6.335	0.985	177.484	182.379	58.002	1779.731
10	5.899	0	178.345	0	58.991	1783.452

Renegotiating agreement condition at moment $t = 1$ requires reconcidering conditions of internal/external stability and choice of abatement commitments. Calculations presented in Table 25 show that though the stable coalition is still composed of four players, $S_1 = S$, Stackelberg equilibrium $(\mathbf{q}^{S_1}, \mathbf{q}^{F_1}) = (\mathbf{5.387}, \mathbf{4.579})$ in the game $\Gamma_1(S_1)$ differs from the restricted over the period $[1, m]$ Stackelberg equilibrium in the game $\Gamma_0(S)$ $(\mathbf{4.455}, \mathbf{3.786})$.

Continuing our reasoning in such a manner, we find that renegotiation takes place at $t = 2, 3, \ldots, m-1$ and causes an increase of abatement commitments, keeping the coalition structure S unchanged. In Fig. 12 step-wise emission reduction under renegotiation in the present example is illustrated. Underlined values correspond to Stackelberg equilibrium in the current game after the renegotiation (*i.e.* realization of the scheme (62)), non-underlined values correspond to the previous game and describe restriction of Stackelberg equilibrium to the remaining period of time.

FIGURE 12 Example 7.1.1 Emission reduction scheme under renegotiation.



7.2 Free-riding from IEA Compliance

Here we analyze time-consistency of an internally stable agreement assuming possible free-riding from IEA compliance. In this case a signatory considers possible deviation from the coalition S and abating less during $[t, t+1)$, knowing that such deviation can be detected at the following stage t and causes exclusion from the coalition. A player decides to defect the agreement during $[t, t+1)$ if its payoff over periods $[t, t+1)$ and $[t+1, m)$ is bigger than its payoff when it follows the agreement. Upon such scenario condition of time-consistency of a self-enforcing coalition receives the following, different in comparison to Definition 5.0.2, presentation.

Definition 7.2.1 *A self-enforcing coalition S, characterized by $\mathbf{n} = (n_1, \dots, n_K)$ of players of K types, is time-consistent under a given abatement scheme, if for every time t, $t = 0, \dots, m-1$, the following conditions hold simultaneously*

1) internal time-consistency

$$\pi_i(\mathbf{q}^S[t,t+1), \mathbf{q}^F[t,t+1)) + \pi_i(\mathbf{q}^S[t+1,m], \mathbf{q}^F[t+1,m]) \qquad (63)$$

$$\geq \pi_i(\mathbf{q}^{S\setminus\{i\}}[t,t+1), \mathbf{q}^{F\cup\{i\}}[t,t+1)) + \pi_i(\mathbf{q}^{S\setminus\{i\}}[t+1,m], \mathbf{q}^{F\cup\{i\}}[t+1,m]), \ \forall i \in S,$$

where $(\mathbf{q}^S[t+1,m], \mathbf{q}^F[t+1,m])$ is restriction over the period $[t+1,m]$ of an optimal solution of the game $\Gamma_0(S)$ and $(\mathbf{q}^{S\setminus\{i\}}[t+1,m], \mathbf{q}^{F\cup\{i\}}[t+1,m])$ is a restriction over the period $[t+1,m]$ of an optimal solution of the game $\Gamma_0(S \setminus \{i\})$;

2) external time-consistency

$$\pi_i(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]) \geq \pi_i(\mathbf{q}^{S\cup\{i\}}[t,m], \mathbf{q}^{F\setminus\{i\}}[t,m]), \quad \forall i \in F,$$

where $(\mathbf{q}^{S\cup\{i\}}[t,m], \mathbf{q}^{F\setminus\{i\}}[t,m])$ is restriction over the period $[t,m]$ of an optimal solution of the game $\Gamma_0(S \cup \{i\})$.

Theorem 7.2.1 *Let condition (28) of environmental efficiency hold. Then internal stability of an agreement, which is self-enforcing in the game $\Gamma_0(S)$ upon Stackelberg equilibrium concept, is time-consistent in terms of Definition 7.2.1 under the abatement scheme (51).*

P r o o f. Let us consider an agreement, which is self-enforcing at $t = 0$. Since free-riders are forbidden to join the agreement after it has been formed, external time-consistency holds. Hence it is necessary to prove that internal stability satisfies conditions of time-consistency (40), that is

$$\pi_i(\mathbf{q}^S[t,t+1), \mathbf{q}^F[t,t+1)) + \pi_i(\mathbf{q}^S[t+1,m], \mathbf{q}^F[t+1,m])$$
$$\geq \pi_i(\mathbf{q}^{S\setminus\{i\}}[t,t+1), \mathbf{q}^{F\cup\{i\}}[t,t+1)) + \pi_i(\mathbf{q}^{S\setminus\{i\}}[t+1,m], \mathbf{q}^{F\cup\{i\}}[t+1,m]),$$

for all $i \in S$ and $t = 1, \dots, m-1$. Net benefit of a signatory of type i, $i = 1, \dots, K$, is

$$\begin{aligned}
&\pi_i(\mathbf{q}^S[t,t+1), \mathbf{q}^F[t,t+1)) \\
&= \frac{b}{N}\left(a(t)Q[t,t+1) - \frac{1}{2}Q^2[t,t+1)\right) - \frac{1}{2}c_i\left(q_i^{S\cup\{i\}}[t,t+1)\right)^2 \\
&= \frac{b}{N}\left((a - \frac{2^t-1}{2^t}Q)\frac{1}{2^{t+1}}Q - \frac{1}{2}\frac{1}{4^{t+1}}Q^2\right) - \frac{1}{2}c_i\frac{1}{4^{t+1}}q_i^S \\
&= \frac{1}{4^{t+1}}\pi_i(\mathbf{q}^S, \mathbf{q}^F) + \frac{2^{t+1}-1}{4^{t+1}}\frac{b}{N}(a-Q)Q + \frac{1}{4^{t+1}}\frac{b}{N}Q^2,
\end{aligned}$$

$$\begin{aligned}
&\pi_i(\mathbf{q}^S[t+1,m], \mathbf{q}^F[t+1,m]) \\
&= \frac{b}{N}\left(a(t+1)Q[t+1,m] - \frac{1}{2}Q^2[t+1,m]\right) - \frac{1}{2}c_i\left(q_i^S[t+1,m]\right)^2, \\
&= \frac{b}{N}\left((a - (1 - \frac{1}{2^{t+1}})Q)\frac{1}{2^{t+1}}Q - \frac{1}{2}\frac{1}{2^{2(t+1)}}Q^2\right) - \frac{1}{2}\frac{1}{2^{2(t+1)}}c_i\left(q_i^S\right)^2 \\
&= \frac{1}{4^{t+1}}\pi_i(\mathbf{q}^S, \mathbf{q}^F) + \frac{2^{t+1}-1}{4^{t+1}}\frac{b}{N}Q(a-Q).
\end{aligned}$$

Lets now assume that a signatory of type i makes a decision to defect the agreement S at a certain moment t. It is assumed that a player knows that the withdrawal will be detected not earlier than at the next moment t starts "cheating" and acts as a free-rider during $[t, t+1)$, reducing $q_i^{F\cup\{i\}}/2^{t+1}$ instead of $q_i^S/2^{t+1}$. Let us denote

$$\tilde{\mathbf{q}}[t,t+1) = \left(\{q_l^S[t,t+1)\}_{l\in S\setminus\{i\}}, \{q_l^F[t,t+1)\}_{l\in F}, q_i^{F\cup\{i\}}[t,t+1)\right)$$

strategies of the players on the current stage. Then its net befit can be presented in the following way

$$\begin{aligned}
&\pi_i(\tilde{\mathbf{q}}[t,t+1)) \\
&= \frac{b}{N}\left(a(t)(Q[t,t+1) - q_i^S[t,t+1) + q_i^{F\cup\{i\}}[t,t+1))\right. \\
&\left. - \frac{1}{2}(Q[t,t+1) - q_i^S[t,t+1) + q_i^{F\cup\{i\}}[t,t+1))^2\right) - \frac{1}{2}c_i\left(q_i^{F\cup\{i\}}[t,t+1)\right)^2 \\
&= \frac{b}{N}\left((a - \frac{2^t-1}{2^t}Q)\frac{1}{2^{t+1}}(Q - q_i^S + q_i^{F\cup\{i\}}) - \frac{1}{2}\frac{1}{4^{t+1}}(Q - q_i^S + q_i^{F\cup\{i\}})^2\right) \\
&- \frac{1}{2}c_i\frac{1}{4^{t+1}}q_i^{F\cup\{i\}} \\
&= \frac{1}{4^{t+1}}\pi_i(\mathbf{q}^{S\setminus\{i\}}, \mathbf{q}^{F\cup\{i\}}) \\
&+ \frac{1}{4^{t+1}}\frac{b}{N}\left(2^{t+1}aQ - 2^{t+1}Q^2 - aQ^{-i} + 2Q^2 - \frac{1}{2}(Q - Q^{-i})\right) \\
&- \frac{1}{4^{t+1}}\frac{b}{N}(q_i^S[t,t+1) - q_i^{F\cup\{i\}}[t,t+1))\left(2^{t+1}a - 2^{t+1} + Q + \frac{1}{2}(q_i^S + q_i^F)\right),
\end{aligned}$$

$$\begin{aligned}
&\pi_i(\mathbf{q}^{S\setminus\{i\}}[t+1,m], \mathbf{q}^{F\cup\{i\}}[t+1,m]) \\
&= \frac{b}{N}\left((a(t+1) - q_i^S[t,t+1) + q_i^{F\cup\{i\}}[t,t+1)))Q^{-i}[t+1,m] - \frac{1}{2}(Q^{-i}[t+1,m])^2\right) \\
&- \frac{1}{2}c_i\left(q_i^{F\cup\{i\}}[t+1,m]\right)^2.
\end{aligned}$$

By the moment $t+1$ the rest of signatories from S recognize the deviator and act further as if coalition structure is $S \setminus \{i\}$, by shifting to the correspondent abatement path $Q^{-i}/2^{t+1}$, which is restriction over $[t+1, m]$ of a Stackelberg solution in the game $\Gamma_0(S \setminus \{i\})$. Then

$$\pi_i(\mathbf{q}^{S\setminus\{i\}}[t+1,m], \mathbf{q}^{F\cup\{i\}}[t+1,m])$$
$$= \frac{1}{4^{t+1}}\pi_i(\mathbf{q}^{S\setminus\{i\}}, \mathbf{q}^{F\cup\{i\}}) + \frac{2^{t+1}-1}{4^t}Q^{-i}(a-Q) + \frac{1}{4^{t+1}}Q^{-i}(q_i^S - q_i^{F\cup\{i\}}).$$

Then

$$(\pi_i(\mathbf{q}^S[t,t+1], \mathbf{q}^F[t,t+1]) + \pi_i(\mathbf{q}^S[t+1,m], \mathbf{q}^F[t+1,m]))$$
$$-(\pi_i(\mathbf{q}^{S\setminus\{i\}}[t,t+1], \mathbf{q}^{F\cup\{i\}}[t,t+1]) + \pi_i(\mathbf{q}^{S\setminus\{i\}}[t+1,m], \mathbf{q}^{F\cup\{i\}}[t+1,m]))$$
$$= \frac{2}{4^{t+1}}\Delta_i^S + \frac{1}{4^{t+1}}\frac{b}{N}(q_i^S - q_i^{F\cup\{i\}})(2^{t+1}(a-Q) + (Q-Q^{-i}) + \frac{1}{2}(q_i^S - q_i^{F\cup\{i\}}))$$
$$+ \frac{1}{4^{t+1}}\frac{b}{N}(Q-Q^{-i})((2^{t+1}-1)(a-Q) - \frac{1}{2}(a-Q) - \frac{1}{2}(a-Q^{-i})) \geq 0.$$

Hence during $[0,m]$ internal time-consistency of a self-enforcing agreement in terms of Definition 7.2.1 holds[1] and no signatory has incentives to defect the agreement. ■

7.3 Potential Time-consistency

General definition of internal time-consistency (see Definition 5.0.2) imposes quite strict conditions on the properties of the agreement, formed as a self-enforcing coalition in the game $\Gamma_0(S)$, so that those requirements can not be held with some coalitions. In this section we introduce a definition of *potential* internal stability of a set of embedded coalitions. This definition proposes transfer mechanisms among signatories to come into action during multistage emission reduction, allowing coalition structure to evolute into another stable condition.

In order to do so, let us first consider a practical generalized notion of coalition S surplus, available during period $[t,m]$,

$$\Delta_i^S(t) = \pi_i\left(\mathbf{q}^S[t,m], \mathbf{q}^F[t,m]\right) - \pi_i\left(\mathbf{q}^{S\setminus\{i\}}[t,m], \mathbf{q}^{F\cup\{i\}}[t,m]\right),$$

where $\Delta_i^S[0,m] = \Delta_i^S$.

Definition 7.3.1 *Let S_0 be a self-enforcing coalition upon Stackelberg solution in the game $\Gamma_0(S_0)$. A set $\{S_0, S_1, S_2, \dots, S_{m-1}\}$ of embedded coalitions ($S_j \subseteq S_{j+1}$, $j = 0, \dots, m-2$) is potentially internally stable if one of the following conditions hold*

a)

$$\sum_{i \in S_t} \Delta_i^{S_t}(t) \geq 0, \quad t = 0, \dots, m-1, \tag{64}$$

1 Here we assume that $(q_i^S - q_i^{F\cup\{i\}}) > 0$, otherwise there is no free-riding from compliance.

b) *if there exists* $\tau \in [1, m-1] : \sum_{i \in S_\tau} \Delta_i^{S_\tau}(\tau) < 0$, *then*

$$\sum_{t=0}^{T} \sum_{i \in S} \Delta_i^{S_t}(t) \geq 0, \tag{65}$$

where T *is the closest to* $m-1$ *integer number such that*

$$\sum_{i \in S_t} \Delta_i^{S_t}(t) \geq 0, t \in [T+1, m-1].$$

Condition (64) has taken place when we were describing potential internal stability of a single coalition in the static game. It requires coalition surplus to be non-negative at every step of the dynamic process, so that it is possible to share $\sum_{i \in S_t} \Delta_i^{S_t}(t)$ among coalition members at every t or to reallocate $\sum_{t=0}^{m-1} \sum_{i \in S_t} \Delta_i^{S_t}(t)$ (which is positive) over all period $[0, m-1]$ according to a certain sharing rule.

Alternative condition (65) is less strict and does not require surplus $\sum_{i \in S_t} \Delta_i^{S_t}(t)$ to be non-negative at every t, but that sum of both positive and negative elements should be non-negative. It means that if $T, T \leqslant m-1$, is the last step, when coalitional surplus is negative, then sum of coalitional surpluses over $[0, T]$ should be non-negative and thus make it possible to provide redistribution of aggregate surplus over the steps. This approach can also be regarded as a *delayed* payment, which implies that we save some part of $\sum_{i \in S_t} \Delta_i^{S_t}(t) \geqslant 0$ unshared at step t, to use it at later steps when surplus is negative.

So far we concentrated our attention on examining time-consistency of the agreement structure and behavior of abatement dynamics. On the other hand little has been said about how renegotiations affect the total amount of reduced emission. The following section is going to answer that question.

7.4 Total Emission Reduction

Let us now evaluate total emission reduction, which can appear under multistage abatement (60) and (61) in the renegotiation game $\Gamma_{t^*}(S)$. The following notation will be practical for further analysis

$$\alpha = \frac{(1-g)^2(\bar{1}, \mathbf{n})((\lambda, \mathbf{n})}{(\bar{1}, \mathbf{N}) + (1-g)^2(\bar{1}, \mathbf{n})(\lambda, \mathbf{n})}, \tag{66}$$

so that with the help of (66) it is possible to write $Q = \alpha a$ and $Q[t, m] = \alpha a(t)$.

Theorem 7.4.1 *In the renegotiation game* $\Gamma_{t^*}(S_{t^*})$ *total emission reduction can only increase.*

P r o o f. Using notation (66) we can represent parameter $a(t)$, which describes environmental situation during renegotiation in the following way

$$a(t) = a - \left(Q - \frac{1}{2^t} Q\right) = a\left(1 - \frac{2^t - 1}{2^t} \alpha\right),$$

where Q is total emission reduction in the game $\Gamma_0(S)$. According to the optimal abatement scheme (51), emission reduction during $[t, m]$ should be

$$Q[t,m] = \frac{1}{2^t}\alpha a,$$

if at step t^* occurs renegotiation, it would impose another abatement target

$$Q^{R_{t^*}} = \alpha a(t) = \alpha a \left(1 - \alpha + \frac{\alpha}{2^t}\right).$$

Let us compare these results

$$Q^{R_{t^*}} - Q[t,m] = \frac{(1-\alpha)(2^t - 1)}{2^t},$$

here $0 \leq \alpha \leq 1$, hence

$$Q^{R_{t^*}} - Q[t,m] \geq 0.$$

Analogous reasoning can be made on further renegotiation steps. ■

Corollary 7.4.1 *Let, in the game $\Gamma_t(S)$, renegotiation on emission reduction targets occur in each step, then total abatement is equal to*

$$Q_{total} = a\left(1 - (1-\alpha)\left(1 - \frac{\alpha}{2}\right)^{m-1}\right); \tag{67}$$

where coefficient α is given in (66), and satisfies condition

$$Q \leq Q_{total} \leq a,$$

where a is initial characteristics of global pollution. Moreover,

$$\lim_{t\to\infty} Q_{total} = a.$$

P r o o f. To calculate total emission reduction, we should summarize abatements achieved over m-steps.

Step 0 announced abatement target is $Q = \alpha a$.

Step 1 emission reduction according to the proposed optimal abatement scheme (51) during $[0, 1)$ is

$$\frac{1}{2}Q = \frac{\alpha}{2}a;$$

current environmental setting is given by

$$a(1) = a - \frac{\alpha}{2}a = a\left(1 - \frac{\alpha}{2}\right);$$

announced renegotiated abatement target is

$$Q[1,m] = \alpha a(1) = \alpha(a - \frac{1}{2}Q) = \alpha\left(1 - \frac{\alpha}{2}\right)a.$$

Step 2 emission reduction during $[1,2)$ is

$$\frac{1}{2}Q[1,m] = \frac{\alpha}{2}\left(1-\frac{\alpha}{2}\right)a;$$

current state of environment is described by parameter $a(2)$

$$a(2) = a(1) - \frac{\alpha}{2}\left(1-\frac{\alpha}{2}\right)a = a\left(1-\frac{\alpha}{2}\right)^2;$$

announced renegotiated abatement target is

$$Q[2,m] = \alpha a(2) = \alpha\left(1-\frac{\alpha}{2}\right)^2 a.$$

step $m-1$ emission reduction during $[m-2, m-1)$ is

$$\frac{1}{2}Q[m-2,m] = \frac{\alpha}{2}\left(1-\frac{\alpha}{2}\right)^{m-2} a;$$

$$\frac{\alpha}{2}\left(1-\frac{\alpha}{2}\right);$$

announced renegotiated abatement target is

$$Q[m-1,m] = \alpha a(m-1) = \alpha\left(1-\frac{\alpha}{2}\right)^{m-1} a.$$

Thus total abatement during $[0,m]$ is

$$\frac{\alpha}{2}a + \frac{\alpha}{2}\left(1-\frac{\alpha}{2}\right)a + \frac{\alpha}{2}\left(1-\frac{\alpha}{2}\right)^2 a + \ldots + \alpha\left(1-\frac{\alpha}{2}\right)^{m-1} a$$

$$= a(1-(1-\alpha)\left(1-\frac{\alpha}{2}\right)^{m-1}). \quad \blacksquare$$

Dynamics of a stock pollutant is related to how countries decide to proceed with emission reduction during the life-cycle of the agreement. In order to model countries' realistic behavior, we constructed the time-consistent abatement scheme. It appeared that such stepwise emission reduction leads to violation of the IEA dynamic stability after a certain threshold moment of time, thus providing countries with incentives to withdraw or access the agreement.

Necessity to handle potential vulnerability of the agreement structure resulted in repeated IEA negotiation, which implied reconsideration of countries' membership status and emission reduction targets. Analysis of the situation revealed that renegotiation eventually reassigned abatement commitments to the agreement members but left the agreement structure unaffected. Moreover, it also appeared that renegotiations would take place with regularity, sufficiently increasing total emission reduction.

8 CONCLUSION

To perform this complex study of IEA formation and performance, we used game theory to explain the strategic possibilities and incentives of countries when negotiation joint efforts on environmental issues such as transboundary pollution. We accomplished this by linking the economic activity of the countries with the physical state of the environment. This link was given by the social welfare function, and expressed via an economic-ecological model of the world. Representing an IEA as a coalitional game of heterogeneous players allowed us to identify the abatement targets of the countries and their agreement membership status. We explored the relationship between the model parameters and cooperation levels, and performed a sensitivity analysis of the agreement's stability. It revealed that large coalitions, including the grand coalition, are only stable if gains from cooperation are small (coalition formation is rather insignificant). If gains from cooperation are large, rather small coalitions are formed.

We specified mechanisms intended to enhance collaboration, including side payments and emission trading, the latter being comprised of setting the price and amount of tradeable pollution permits, and the initial reallocation of abatement commitments. It was shown that such initiatives positively effect the size and structure of a potentially stable coalition, making agreements with higher abatement targets possible.

Next, we examined the dynamics of the agreement, and addressed the question that if both members and free-riders have incentives to change their status during the agreement life-cycle, is there a threat to compliance with the agreement? The intuition behind this question comes from the idea that the dynamics of the pollution flow puts continued agreement stability in jeopardy.

Pollution flow dynamics, and thus the potential for long-term multilateral collaboration, is related to how countries perform emission reduction during the life-cycle of the agreement. In order to realistically model countries' behavior, we suggest an algorithm for constructing an abatement scheme. The scheme assigns emission reductions so that the abatement effort during each time period is adjusted according to the pollution reduction undertaken during the previous stage, and takes into account any change in the environment.

Following this scheme, internal dynamic stability of the IEA is achieved with respect to free-riding in compliance and participation. However, external dynamic stability is violated after a certain threshold moment in time, possibly providing non-signatories with incentives to access the agreement and signatories with incentives to leave. A renegotiation is required to stabilize the agreement, in which membership status and emission reduction targets are reconsidered. We showed that the renegotiation process enhances an IEA's dynamic stability.

Finally, we examined compliance with the abatement targets and the effect of the agreement on the global level of pollution and the welfare of countries. This analysis revealed that following the time-consistent scheme of emission reduction would prevent free-riding from compliance, and in case of renegotiation, the abatement efforts of the players would only increase.

There are many possibilities for the extension of this research, including the following:

1. Softening the assumption of perfect information and considering incomplete information, where players of one type do not have information about the benefit functions of players of another type, [Kolstad 2003], [Wirl 2004], [Courtois & Haeringer 2005],[Kryazhimskii *et al.* 1998].

2. Calibration of the model parameters, which has a crucial role in applications of this research. Using synthetic and real world data, one should observe the relationship between input and output values to make estimations of the model parameters.

3. The formation of multiple coalitions resulting in multiple IEAs. One could study the coalition formation process as in [Finus 2001], [Alcalde & Revilla 2001], [Cesar & De Zeeuw 1996], [Folmer *et al.* 1993].

4. The combination of the two major aspects of free-riding, namely free-riding from membership and free-riding from compliance.

5. The study of methods to increase the size of a self-enforcing coalition.

The methodological approach, as well as the theoretical and numerical analysis, presented in this thesis provides a formidable framework for future study of the theory of IEA formation and sustainability.

APPENDIX 1

FIGURE 13 Example 3.2.1 Data Set.

n1	n2	q_1^1	q_1^F	π_1^S	π_1^F	Q	Π
0	0	0	5,333333333	0	272,8888889	60	4120
0	1	2,325130789	5,502809533	275,2050636	268,9863937	58,7289285	4069,402931
0	2	4,829945663	5,393439324	272,9590125	271,5188094	59,54920507	4102,933661
0	3	7,130124777	4,991087344	273,9156269	280,3975585	62,56684492	4209,871601
0	4	8,887243103	4,369561192	277,7880639	292,7605702	67,22829106	4349,418276
0	5	9,958506224	0	283,5408022	0	72,61410788	4480,208674
1	0	0	5,439624853	0	270,4556064	59,2028136	4088,405998
1	1	4,722317871	5,391312903	273,259034	271,5675419	59,56515323	4103,511416
1	2	7,051564566	5,053621272	273,0164027	279,0627714	62,09784046	4194,012748
1	3	8,888477385	4,481274015	275,9286693	290,6594726	66,39044489	4325,319631
1	4	10,0548829	3,787339227	281,1632813	302,8524642	71,5949558	4455,921934
1	5	10,5785124	0	287,5076839	0	76,85950413	4561,505362
2	0	0	5,384615385	0	271,7209073	59,61538462	4105,399408
2	1	6,968641115	5,110336818	272,2262016	277,8378058	61,67247387	4179,424298
2	2	8,8823094	4,58919319	274,1206768	288,5793583	65,58105107	4301,44406
2	3	10,14541765	3,922894826	278,7463888	300,6314877	70,57828881	4430,88793
2	4	10,75268817	3,225806452	284,9173315	311,2209504	75,80645161	4539,97572
2	5	10,83870968	0	291,4769337	0	80,64516129	4620,811655
3	0	0	5,161290323	0	276,725633	61,29032258	4166,146375
3	1	8,86876992	4,693057416	272,373086	286,530659	64,80206938	4277,905102
3	2	10,22974298	4,057798048	276,2981886	298,3436677	69,56651464	4405,165401
3	3	10,92523045	3,368612723	282,2165088	309,219786	74,73540457	4517,294144
3	4	11,07120498	2,715081222	288,8685646	317,668543	79,63689084	4602,739209
3	5	10,8398024	0	295,3280452	0	83,90966831	4663,561083
4	0	0	4,792626728	0	284,5236326	64,05529954	4254,81535
4	1	10,3075073	4,191719636	273,8274438	295,9959921	68,56210273	4378,822319
4	2	11,09570042	3,513638465	279,4065878	307,0988159	73,64771151	4493,441264
4	3	11,30703998	2,853681518	286,1019859	316,0283996	78,59738861	4583,465249
4	4	11,10982525	2,268255989	292,8294317	322,4002397	82,98808008	4648,616768
4	5	10,66666667	0	298,962963	0	86,66666667	4694,814815
5	0	0	4,324324324	0	293,596299	67,56756757	4351,93572
5	1	11,26363956	3,660682858	276,4898162	304,8570605	72,54487856	4468,404086
5	2	11,54599498	2,996460603	283,170628	314,253434	77,52654548	4562,919732
5	3	11,38756265	2,396132974	290,1489628	321,1327452	82,0290027	4632,598123
5	4	10,95640265	1,886936012	296,6466545	325,7672124	85,84797991	4682,362205
5	5	10,38399135	0	302,3189514	0	88,96700919	4718,421751
6	0	0	3,80952381	0	302,4943311	71,42857143	4442,176871
6	1	11,78781925	3,143418468	280,0681383	312,3360391	76,42436149	4541,031569
6	2	11,67315175	2,529182879	287,2737791	319,7402055	81,0311284	4615,418099
6	3	11,25703565	2,001250782	294,1437366	324,8226984	84,99061914	4669,012176
6	4	10,67995728	1,566393734	300,2174815	328,119456	88,25204699	4707,851777
6	5	10,03802281	0	305,3670479	0	90,87452471	4737,02092
7	0	0	3,294538944	0	310,2686113	75,29095792	4517,729178
7	1	11,96671757	2,667580791	284,1902749	318,2118604	79,99314407	4596,983122
7	2	11,56905278	2,12099301	291,4376915	323,7737842	84,09255242	4654,68162
7	3	10,98875941	1,666628511	297,9370301	327,4308258	87,50028617	4696,549691
7	4	10,33086149	1,299184097	303,4868848	329,7465898	90,25611927	4727,856205
7	5	9,660377358	0	308,1039516	0	92,45283019	4752,319924
8	0	0	2,811501597	0	316,5361832	78,91373802	4577,192785
8	1	11,89296333	2,246448629	288,5104427	322,609454	83,15163528	4639,278865
8	2	11,31115091	1,772080309	295,4597959	326,6602625	86,70939768	4684,444674
8	3	10,63744057	1,386090741	301,4422087	329,2506823	89,60431945	4718,093616
8	4	9,944064636	1,077273669	306,4362557	330,8672314	91,92044748	4744,217676
8	5	9,27191679	0	310,5430064	0	93,75928678	4765,389277
9	0	0	2,377919321	0	321,3175202	82,1656051	4622,702747
9	1	11,64794098	1,883083792	292,7659425	325,7980736	85,87687156	4671,457615
9	2	10,95867666	1,477760943	299,2156022	328,6928063	88,91679293	4707,675705
9	3	10,24163866	1,152184349	304,6214263	330,5123347	91,35861738	4735,620493
9	4	9,542889477	0,8931165792	309,0710411	331,6383117	93,30162566	4758,1125
9	5	8,885950413	0	312,7068078	0	94,84297521	4776,890868
10	0	0	2	0	324,8333333	85	4657,5
10	1	11,29556735	1,57453363	296,7874551	328,0651265	88,19099778	4696,537989
10	2	10,55408971	1,231310466	302,6433957	330,1115667	90,7651715	4726,482457
10	3	9,827390702	0,9575406325	307,4697737	331,3849547	92,81844526	4750,418475
10	4	9,142359599	0,7400957771	311,4106328	332,1693821	94,44928167	4770,277059
10	5	8,510638298	0	314,6220009	0	95,74468085	4787,234043

FIGURE 14 Example 3.2.1 Data Set.

n1	n2	q_2^S	q_2^F	π_2^S	π_2^F	Q	Π
0	0	0	3,333333333	0	275,5555556	60	4120
0	1	1,406799531	3,399765533	277,0613521	273,2296238	59,2028136	4088,405998
0	2	2,884615385	3,365384615	275,6410256	274,4391026	59,61538462	4105,399408
0	3	4,301075269	3,225806452	275,9856631	279,2230316	61,29032258	4166,146375
0	4	5,529953917	2,995391705	278,0337942	286,6770017	64,05529954	4254,81535
0	5	6,486486486	2,702702703	281,4414414	295,3494035	67,56756757	4351,93572
0	6	7,142857143	2,380952381	285,7142857	303,8548753	71,42857143	4442,176871
0	7	7,520143241	2,05908684	290,3610862	311,2861725	75,29095792	4517,729178
0	8	7,667731629	1,757188498	294,9946752	317,277234	78,91373802	4577,192785
0	9	7,643312102	1,486199575	299,3630573	321,8476296	82,1656051	4622,702747
0	10	7,5	0	303,3333333	0	85	4657,5
1	0	0	3,439255958	0	271,8252293	58,7289285	4069,402931
1	1	2,951448669	3,369570564	275,3496859	274,2925032	59,56515323	4103,511416
1	2	4,355400697	3,193960511	276,7788853	280,2861379	61,67247387	4179,424298
1	3	5,5429812	2,933160885	279,7469998	288,5954829	64,80206938	4277,905102
1	4	6,442192063	2,619824772	283,7878851	297,6432278	68,56210273	4378,822319
1	5	7,039774727	2,287926786	288,3838389	306,1133666	72,54487856	4468,404086
1	6	7,367387033	1,964636542	293,0949523	313,2623903	76,42436149	4541,031569
1	7	7,479198479	1,667237994	297,6154932	318,8789842	79,99314407	4596,983122
1	8	7,433102081	1,404030393	301,7706843	323,0825663	83,15163528	4639,278865
1	9	7,279963115	1,17692737	305,4854296	326,1305116	85,87687156	4671,457615
1	10	7,059729591	0	308,7490028	0	88,19099778	4696,537989
2	0	0	3,370899577	0	274,2459207	59,54920507	4102,933661
2	1	4,407227854	3,158513295	277,6780805	281,4570609	62,09784046	4194,012748
2	2	5,551443375	2,868245744	281,5171224	290,5537984	65,58105107	4301,44406
2	3	6,393589361	2,53612378	286,108905	299,8873294	69,56651464	4405,165401
2	4	6,93481276	2,196024041	290,9485785	308,2562211	73,64771151	4493,441264
2	5	7,216246864	1,872787877	295,6684405	315,0951943	77,52654548	4562,919732
2	6	7,295719844	1,5807393	300,0483858	320,3399023	81,0311284	4615,418099
2	7	7,23065799	1,325620631	303,9854711	324,195529	84,09255242	4654,68162
2	8	7,069469319	1,107550193	307,4543711	326,9546627	86,70939768	4684,444674
2	9	6,84917291	0,9236005894	310,4742829	328,8975355	88,91679293	4707,675705
2	10	6,596306069	0	313,0860966	0	90,7651715	4726,482457
3	0	0	3,11942959	0	282,7329603	62,56684492	4209,871601
3	1	5,555298366	2,800796259	283,3353909	292,542143	66,39044489	4325,319631
3	2	6,340886033	2,451809266	288,3960294	302,0742162	70,57828881	4430,88793
3	3	6,828269034	2,105382952	293,4065707	310,283619	74,73540457	4517,294144
3	4	7,066899987	1,783550949	298,087844	316,7918526	78,59738861	4583,465249
3	5	7,117226654	1,497583108	302,3061424	321,6710065	82,0290027	4632,598123
3	6	7,03564728	1,250781739	306,0238165	325,1981676	84,99061914	4669,012176
3	7	6,867974634	1,04164282	309,2576082	327,6912306	87,50028617	4696,549691
3	8	6,648400355	0,8663067129	312,0505033	329,4307993	89,60431945	4718,093616
3	9	6,401024164	0,7201152184	314,4549727	330,6367905	91,35861738	4735,620493
3	10	6,142119189	0	316,5239245	0	92,81844526	4750,418475
4	0	0	2,730975745	0	294,550545	67,22829106	4349,418276
4	1	6,284301814	2,367087017	290,6414691	304,1972084	71,5949558	4455,921934
4	2	6,720430108	2,016129032	295,7567349	312,1964967	75,80645161	4539,97572
4	3	6,919503114	1,696925764	300,3596502	318,3596367	79,63689084	4602,739209
4	4	6,943640782	1,417659993	304,4008271	322,8825821	82,98808008	4648,616768
4	5	6,847751655	1,179335007	307,9006632	326,1010118	85,84797991	4682,362205
4	6	6,6749733	0,978996084	310,9107459	328,34948	88,25204699	4707,851777
4	7	6,456788433	0,8119900605	313,4925129	329,9048285	90,25611927	4727,856205
4	8	6,215040398	0,6732960431	315,7066702	330,97603	91,92044748	4744,217676
4	9	5,964305923	0,558197862	317,608548	331,7130921	93,30162566	4758,1125
4	10	5,71397475	0	319,2465146	0	94,44928167	4770,277059
5	0	0	2,282157676	0	306,2504663	72,61410788	4480,208674
5	1	6,611570248	1,928374656	297,9987706	313,9964635	76,85950413	4561,505362
5	2	6,774193548	1,612903226	302,4904613	319,8057579	80,64516129	4620,811655
5	3	6,7748765	1,340860974	306,3437936	323,9842109	83,90966831	4663,561083
5	4	6,666666667	1,111111111	309,6296296	326,9135802	86,66666667	4694,814815
5	5	6,489994592	0,9194159005	312,4277585	328,9376402	88,96700919	4718,421751
5	6	6,273764259	0,7604562738	314,8134762	330,3262059	90,87452471	4737,02092
5	7	6,037735849	0,6289308176	316,8529726	331,2764527	92,45283019	4752,319924
5	8	5,794947994	0,5200594354	318,6025477	331,9269319	93,75928678	4765,389277
5	9	5,553719008	0,4297520661	320,1093186	332,3729618	94,84297521	4776,890868
5	10	5,319148936	0	321,4124038	0	95,74468085	4787,234043

FIGURE 15   Example 3.2.1 Numerical Simulations with Two Types of Players.

b		c2=0.01 c1=p*c2				c2=0.1 c1=p*c2			
		Asymmetry				Asymmetry			
		Hi(p=0,1)		Lo(p=0,8)		Hi(p=0,1)		Lo (p=0,8)	
	0.01	n1=0,n2=7	Q=85 Π=49	n1=0,n2=3	Q=54 Π=39	n1=0,n2=3	Q=36 Π=29	n1=0,n2=2	Q=11 Π=10
				n1=1,n2=2					
				n1=2,n2=1					
				n1=3,n2=0					
	0.1	n1=7,n2=N2	Q=98 Π=499	n1=0,n2=9	Q=92 Π=495	n1=0,n2=7	Q=85 Π=485	n1=0,n2=3	Q=55 Π=390
		n1=8,n2=7		n1=1,n2=8				n1=1,n2=2	
		n1=9,n2=3		n1=2,n2=7				n1=2,n2=1	
		n1=N1,n2=7		n1=4,n2=4				n1=3,n2=0	
				n1=5,n2=3					
				n1=6,n2=2					
				n1=7,n2=1					
				n1=8,n2=0					
	1	n1=N1,n2=7	Q=100 Π=5000	n1=8,n2=N2	Q=99 Π=4998	n1=7,n2=N2	Q=98 Π=4994	n1=0,n2=9	Q=92 Π=4951
				n1=9,n2=9		n1=8,n2=7		n1=1,n2=8	
				n1=N1,n2=7		n1=9,n2=3		n1=2,n2=7	
						n1=N1,n2=0		n1=4,n2=4	
								n1=5,n2=3	
								n1=6,n2=2	
								n1=7,n2=1	
								n1=8,n2=0	
	10	**n1=N1,n2=N2**	Q=100 Π=50000	**n1=N1,n2=N2**	Q=100 Π=50000	n1=N1,n2=7	Q=100 Π=50000	n1=8,n2=N2	Q=99 Π=49998
								n1=9,n2=9	
								n1=N1,n2=7	

FIGURE 16 Example 3.2.1 Numerical Simulations with Two Types of Players.

b		c2=1 c1=p*c2				c2=10 c1=p*c2			
		Asymmetry				Asymmetry			
		Hi p=0,1		Lo p=0,8		Hi p=0,1		Lo (p=0,8)	
	0.01	n1=0,n2=2	Q=5 Π=5	n1=0,n2=2	Q=1 Π=1	n1=0,n2=2	Q=0,6 Π=0,5	n1=0,n2=2	Q=0,1 Π=0,1
	0.1	n1=0,n2=3	Q=36 Π=290	n1=0,n2=2	Q=11 Π=100	n1=0,n2=2	Q=5 Π=50	n1=0,n2=2	Q=1 Π=12
	1	n1=0,n2=7	Q=85 Π=4850	n1=0,n2=3	Q=55 Π=3902	n1=0,n2=3	Q=36 Π=2896	n1=0,n2=2	Q=11 Π=992
				n1=1,n2=2					
				n1=2,n2=1					
				n1=3,n2=0					
	10	n1=7,n2=N2	Q=98 Π=49942	n1=0,n2=9	Q=92 Π=49513	n1=0,n2=7	Q=85 Π=48507	n1=0,n2=3	Q=55 Π=39015
		n1=8,n2=7		n1=1,n2=8				n1=1,n2=2	
		n1=9,n2=3		n1=2,n2=7				n1=2,n2=1	
		n1=N1,n2=0		n1=4,n2=4				n1=3,n2=0	
				n1=5,n2=3					
				n1=6,n2=2					
				n1=7,n2=1					
				n1=8,n2=0					

REFERENCES

[Alcalde & Revilla 2001] J. Alcalde, P. Revilla (2001). *Tops responsiveness, strategy-proofness and coalition formation problems*. Working Papers. Serie AD 2001-11, Instituto Valenciano de Investigaciones Económicas, S.A. (Ivie).

[Altamirano-Cabrera & Finus 2006] J.-C. Altamirano-Cabrera, M. Finus (2006). *Permit Trading and Stability of International Climate Agreements*. Journal of Applied Economics **9**, pp. 19–47.

[Asheim *et al.* 2006] G.B.Asheim, C. Bretteville Froyn, J. Hovi, F.C. Menz (2006). *Regional versus global cooperation for climate control*. Journal of Environmental Economics and Management **51**, pp. 93–109.

[d'Aspremont *et al.* 1983] C. d'Aspremont, A. Jacquemin, J. A. Weymark (1983). *On the Stability of Collusive Price Leadership*. Canadian Journal of economics **16**, pp. 17–25.

[d'Aspremont *et al.* 1986] C. d'Aspremont, J. J. Gabszewicz (1986). *On the Stability of Collusion*. (Eds. G.F. Matthewson, J.E. Stiglitz) New Developments in the Analysis of Market Structure, NY, Macmillian, pp. 243–264.

[Aumann 1959] R. J. Aumann (1959). *Acceptable Points in General Cooperative n-Person Games*. (Eds. A. W. Tucker,R. D. Luce) Contributions to the Theory of Games IV, Annals of Mathematics Study **40**, Princeton University Press, Princeton, NJ, pp. 287–324.

[Barrett 1991] S. Barrett (1991). *The Paradox of International Environmental Agreements*. London: London Business School, mimeo.

[Barrett 1992] S. Barrett (1992). *Cooperation and Competition in International Environmental Protection*. Invited paper, Economics of the Environment Session, International Economic Association Meeting, 24-28 August, Moscow.

[Barrett 1994a] S. Barrett (1994). *Self-Enforcing International Environmental Agreements* Oxford Economic Papers **46**, pp. 878–894.

[Barrett 1994b] S. Barrett (1994). *The Biodiversity Supergame*. Environmental and Resource Economics **4**, pp. 111-122.

[Barrett 1997a] S. Barrett (1997). *Toward a Theory of International Environmental Cooperation*. (Eds. C. Carraro) New Directions in the Economic Theory of the Environment, Cambridge: Cambridge University Press, pp. 239–280.

[Barrett 1997b] S. Barrett (1997). *The Strategy of Trade Sanctions in International Environmental Agreements*. Resource and Environmental Economics **19**, pp. 345–361.

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

[Barrett 2001] S. Barrett (2001). *International Cooperation for Sale*. European Economic Review **45**, pp. 1835–1850.

[Barrett 2003] S. Barrett (2003). *Environment and Statecraft The Strategy of Environmental Treaty-Making*. Oxford University Press, 456 pp.

[Bauer 1992] A. Bauer (1992). *International Cooperation over Greenhouse Gas Abatement*. Seminar für empirische Wirtschaftsforschung, University of Munich, mineo.

[Bellman 2003] R. Bellman (2003). *Dynamic programming*. Dover Publications Inc.; Dover Ed edition (27 Jan 2003), 366 p.

[Bernheim *et el.* 1987] B. D. Bernheim, B. Peleg, M. D. Whinston (1987). *Coalition-Proof Nash Equilibria I. Concepts*. Journal of Economic Theory **42**, no. 1, pp. 1–12.

[Bloch 1995] F. Bloch (1995). *Endogenous structures of Association in Oligopolies*. RAND Journal of Economics **26**, pp. 537–556.

[Bloch 1996] F. Bloch (1996). *Sequential Formation of Coalitions with Fixed Payoff Division and Externalities*. Games and Economic Behavior **14**, pp. 90–123.

[Bloch 1997] F. Bloch (1997). *Noncooperative models of coalition formation in games with spillovers*. (Eds. C. Carraro, D. Siniscalco) New Direction in Economic Theory of Environment, Cambridge University Press, Cambridge, pp. 311–352.

[Botteon & Carraro 1997] M. Botteon, C. Carraro (1997). *Burden-sharing and Coalition Stability in Environmental Negotiations with Asymmetric Countries*. (Eds. C. Carraro) International Environmental Negotiations: Strategic Policy Issues, Edvar Elgar, pp. 26–55.

[Botteon & Carraro 1998] M. Botteon, C. Carraro (1998). *Strategies for Environmental Negotiations: Issue Linkage with Heterogeneous Countries*. (Eds. N. Hanley, H. Folmer) Game Theory and the Environment. E. Elgar, Cheltenham, pp. 181–203.

[Bosello *et al.* 2003] F. Bosello, B. Buchner, C. Carraro (2003). *Equity, Development and Climate Change Control*. Journal of the European Economic Association **1**, no. 2-3, pp. 600–611.

[Brechet *et al.* 2007] T. Brechet, F. Gerard, H. Tulkens (2007). *Climate coalitions: a theoretical and computational appraisal*. Université catholique de Louvain, Département des Sciences Economiques Working Paper 2007006, revised.

[Breton *et al.* 2008a] M. Breton, A. Sokri, G. Zaccour (2008). *Incentive equilibrium in an overlapping-generations environmental game*. European Journal of Operational Research **185**, pp. 687–699.

[Breton *et al.* 2008b] M. Breton, L. Sbragia, G. Zaccour (2008). *Dynamic Models for International Environmental Agreements*. Journal of Environmental Economics and Management.

[Le Breton *et al.* 2007] M. Le Breton, V. Makarov, A. Savvateev, S. Weber (2007). *Multiple membership and federal structures*. IDEI Working Paper No. 491.

[Le Breton & Soubeyran 1997] M. Le Breton, A. Soubeyran (1997). *The Interaction between International Environmental and Trade Policies*. (Eds. C. Carraro) International Environmental Negotiations and Strategic Policy Issues, E. Elgar, Cheltenham, pp. 126–149.

[Carraro 2000] C. Carraro (2000). *The Economics of Coalition Formation*. (Eds. J. Gupta, M. Grubb) Climate Change and European Leadership, Kluwer Academic Publishers, pp. 135–156.

[Carraro *et al.* 2006] C. Carraro, J. Eyckmans, M. Finus (2006). *Optimal transfers and participation decisions in international environmental agreements*. Review of International Organizations **1**, pp. 379–396.

[Carraro & Marchiory 2003] C. Carraro, C. Marchiori (2003). *Stable Coalitions*. (Eds. C. Carraro) The Endogenous Formation of Economic Coalitions, pp. 156–198.

[Carraro & Marchiory 2004] C. Carraro, C. Marchiori (2004). *Endogenous Strategic Issue Linkage in International Negotiations*. (Eds. C. Carraro, V. Fragnelli) Game Practice and the Environment.

[Carraro *et al.* 2003] C. Carraro, C. Marchiori, S. Orefice (2003). *Endogeneous Minimum Participation in International Environmental Treaties*. Working paper No. 113.2003 FEEM.

[Carraro & Siniscalco 1993] C. Carraro, D. Siniscalco (1993) Strategies for the International Protection of the Environment, Journal of Public Economics **52**, no. 3, pp. 309–328.

[Carraro & Siniscalco 1998] C. Carraro, D. Siniscalco (1998). *International Environmental Agreements: Incentives and Political Economy*. European Economic Review **42**, pp. 561–572.

[Cesar & De Zeeuw 1996] H. Cesar, A. De Zeeuw (1996). *Issue Linkage in Global Environmental Problems*. (Eds. A. Xepapadeas) Economic Policy for the Environment and Natural Resources. Cheltenham: Edward Elgar.

[Chandler & Tulkens 1995] P. Chander, H. Tulkens (1995). *A Core-Theoretic Solution fo the Design of Cooperative Agreements on Transfrontier Pollution*. International Tax and Public Finance **2**, pp. 279–293.

[Chandler & Tulkens 1997] P. Chander, H. Tulkens (1997). *The Core of an Economy with Multilateral Environmental Externalities*. International Journal of Game Theory **26**, no. 3, pp. 379–401.

[Chwe 1994] M. Chwe (1994). *Firsighted Coalition Stability*. Journal of Economic Theory **63**, pp. 299–325.

[Copeland & Taylor 2003] B. R. Copeland, M. S. Taylor (2003). *Trade and Environment: Theory and Evidence*. Princeton University Press, 304 p.

[Courtois & Haeringer 2005] P. Courtois, G. Haeringer (2005). *The Making of International Environmental Agreements*, Working Paper.

[Dementieva 2004] M. Dementieva (2004). *Regularization in Multistage Cooperative Games*. Jyväskylä Series in Computing 42. Available on-line `http://dissertations.jyu.fi/studcomp/9513919293.pdf`

[Demetieva & Pavlova 2007a] M. Dementieva, Yu. Pavlova (2007). *Self-Enforcing International Environmental Agreement in the Game with Heterogeneous Players*. Reports of the Department of Mathematical Information Technology Series B. Scientific Computing, B 1/2007.

[Dementieva & Pavlova 2007b] M. Dementieva, Yu. Pavlova (2007). *Emission Trading as an Expanding Mechanism of International Environmental Agreement*. (Eds. P. Neittaanmäki, J. Périaux, T. Tuovinen) Evolutionary and Deterministic Methods for Design, Optimization and Control, CIMNE, Barcelona, Spain.

[Dementieva *et al.* 2007] M. Dementieva, Yu. Pavlova, V. Zakharov (2007). *Dynamic Regularization of Self-Enforcing International Environmental Agreement in the Game of Heterogeneous Players*. (Eds. L. Petrosjan, N. Zenkevich) International Conference "Game Theory and Management" GTM2007, St. Petersburg, Russia, 28-29 June 2007, Collected Papers, pp. 68–92.

[Dementieva *et al.* 2008] M. Dementieva, Yu. Pavlova, V. Zakharov (2008). *Dynamic Regularization of Self-Enforcing International Environmental Agreement in the Game of Heterogeneous Players*. (Eds. L. Petrosjan, V. Mazalov) Game Theory and Applications 14, pp. 2–20.

[Diamantoudi & Sartzetakis 2006] E. Diamantoudi, E. S. Sartzetakis (2006). *Stable International Environmental Agreements: An Analytical Approach*. Journal of Public Economic Theory **8**, no. 2, pp. 247–263.

[Eyckmans 2001] J. Eyckmans (2001). *On the Farsighted Stability of the Kyoto Protocol*. CLIMNEG Working Paper 40, CORE, Universite Catholique de Louvain.

[Eyckmans 2003] J. Eyckmans (2003). *On the Farsighted Stability of International Climate Agreements - the Role of Foresight*. mimeo.

[Eyckmans & Finus 2004] J. Eyckmans, M. Finus (2004). *An Almost Ideal Sharing Scheme for Coalition Games with Externalities*. Katholieke Universiteit Leuven, Center for Economic Studies, Working Paper 2004-14.

[Eyckmans & Tulkens 2003] J. Eyckmans, H. Tulkens (2003). *Simulating Coalitionally Stable Burden Sharing Agreements for the Climate Change Problem* Resource and Energy Economics **25**, no. 4, pp. 299–327.

[Finus 2001] M. Finus (2001). *Game Theory and International Environmental Cooperation (New Horizons in Environmental Economics)*. Edward Elgar Publishing Ltd., 416 pp.

[Finus 2003a] M. Finus (2003). *Stability and Design of International Environmental Agreements: The Case of Global and Transboundary Pollution* (Eds. H. Folmer, T. Tietenberg) International Yearbook of Environmental and Resource Economics, 2003/4, Edward Elgar, Cheltenham, UK, pp. 82–158.

[Finus 2003b] M. Finus (2003). *New developments in coalition theory: an application to the case of global pollution*. (Eds. L. Marsiliani, M. Rauscher, C. Withagen), Environmental policy in an international perspective, pp. 19–49.

[Finus 2004] M. Finus (2004). *Modesty Pays: Sometimes*. FEEM, Nota di lavoro 68.2004.

[Finus 2007] M. Finus (2007). *Game Theoretic Research on the Design og International Environmental Agreements: Insights, Critical Remarks and Future Challenges*. University of Hagen, mimeo.

[Finus & Rundshagen 1998] M. Finus, B. Rundshagen (1998). *Renegotiation-proof equilibria in global emission game when players are impatient*. Environmental and Resource Economics **12**, pp. 275–306.

[Finus & Rundshagen 2000] M. Finus, B. Rundshagen (2000). *Strategic Links between Environmental and Trade Policies if Plant Location Is Endogenous*. University of Hagen, Working Paper No. 283.

[Finus & Rundshagen 2003] M. Finus, B. Rundshagen (2003). *Endogenous Coalition Formation in Global Pollution Control: A Partition Function Approach*. (Eds. C. Carraro) Endogeneous Formation of Economic Coalitions, pp. 199–243.

[Finus & Rundshagen 2006] M. Finus, B. Rundshagen (2006). *Participation in International Environmental Agreements: The Role of Timing and Regulation* Natural Resource Modeling **19**, pp. 165–200.

[Folmer *et al.* 1993] H. Folmer, P. von Mouche, S. Ragland (1993). *Interconnected games and international environmental problems*. Journal Environmental and Resource Economics **3**, no. 4, pp. 313–335.

[Germain *et al.* 2003] M. Germain, P. Toint, H. Tulkens, A. de Zeeuw (2003). *Transfers to sustain dynamic core-theoric cooperation in international stock pollutant control* Journal of Economic Dynamics and Control **28**, pp. 79–99.

[Golombek & Hoel 2005] R. Golombek, M. Hoel (2005). *Climate Policy under Technology Spillovers*. Environmental and Resource Economics **31**, no. 2, pp. 201–227.

[Grübler & Gritsevskyi 2002] A. Grübler, A. Gritsevskyi (2002). *A Model of Endogenous Technological Change through Uncertain Returns on Innovation*. (Eds. A. Grübler *et al.*) Technological Change and the Environment, Resources For the Future, Washington, DC, USA, International Institute for Applied System Analysis, Laxenburg, Austria.

[Hanley *et al.* 1997] N. Hanley, J. F. Shogren, B. White (1997). *Environmental Economics in Theory and Practice*. Macmillian Press LTD, 464 p.

[Hoel 1992] M. Hoel (1992). *International Environmental Conventions: the Case of Uniform Reductions of Emissions*. Environmental and Resource Economics **2**, pp. 141–159.

[Hoel & Schneider 1997] M. Hoel, K. Schneider (1997). *Incentives to Participate in an International Environmental Agreement*. Environmental and Resource Economics **9**, pp. 153–170.

[Ioannidis 2000] A. Ioannidis, A. Papandreou, E. Sartzetakis (2000). *International Environmental Agreements: a Literature Review*. Working Papers, GREEN.

[Jeppensen & Andersen 1998] T. Jeppensen, P. Andersen (1998). *Commitment and Fairness in Environmental Games*. (Eds. N. Hanley, H. Folmer) Game Theory and the Environment, E. Elgar, Cheltenham, pp. 65–83.

[Katsoulacos 1997] Y. Katsoulacos (1997). *R&D Spillover, Cooperation, Subsidies and International Agreements*. (Eds. C. Carraro) International Environmental Negotiations: Strategic Policy Issues, E. Elgar, Cheltenham, pp. 97–109.

[Kolstad 2003] C. D. Kolstad (2007). *Systematic uncertainty in self-enforcing international environmental agreements*. Journal of Environmental Economics and Management **53**, pp. 68–79.

[Kronbak & Lindroos] L. G. Kronbak, M. Lindroos (2007). *Sharing Rules and Stability in Coalition Games with Externalities*. Marine Resource Economics **22**, pp. 137–154.

[Kryazhimskii *et al.* 1998] A.V. Kryazhimskii, A. Nentjes, S. Shibayev, A.M. Tarasyev (1998). *Searching Market Equilibria under Uncertain Utilities*. IIASA Interim Report IR-98-007.

[McGinty 2006] M. McGinty (2006). *International Environmental Agreements among Asymmetric Nations.*, Oxford Economic Papers **59**, no. 1, pp. 45–62.

[Pavlova 2008] Yu. Pavlova (2008). *Dynamic game of international environmental agreement*. Vestnik St. Petersburg University, ser. 10, n 3, pp. 85–97.

[Pavlova 2008] Yu. Pavlova (2008). *Free-riding Aspects in Dynamic Game-Theoretic Model of International Environmental Agreements*. 13th International Symposium on Dynamic Games and Applications, Wroclaw, Poland, 30 June - 3 July 2008, CD of Full Papers, submitted to Annals of International Society of Dynamic Games **11**.

[Pavlova *et al.* 2008] Yu. Pavlova, M. Dementieva, V. Zakharov (2008). *Role of Renegotiation for Time-Consistency of a Self-Enforcing International Environmental Agreement*. 2nd International Conference on Game Theory and Management, St. Petersburg, Russia, 26-27 June 2008, Book of Abstracts, pp. 213–214.

[Pavlova *et al.* 2008] Yu. Pavlova, M. Dementieva, V. Zakharov (2008). *A Multistage Game-Theoretic Model of International Environmental Agreements: Stability and Environmental Efficiency*. 13th International Symposium on Dynamic Games and Applications, Wroclaw, Poland, 30 June - 3 July 2008, CD of Full Papers (submitted to JEEM)

[Petrosjan 1977] L. Petrosjan (1977). *Stability of solutions in n-person differential games*. Bulletin of Leningrad University **19**, pp. 46–52. (Russian)

[Petrosjan & Danilov 1979] L. Petrosjan, N. Danilov (1979. *) Stability of the solutions in nonantagonistic differential games with transferable payoffs*. (Russian) Vestnik Leningrad. Univ. Mat. Mekh. Astronom. vyp. 1, 52–59, 134. (Reviewer: A. B. Kurzhanskiui)

[Petrosjan & Zaccour 2000] Petrosjan L., Zaccour G. (2000). *A multistage supergame of downstream pollution*. Annals of International Society of Dynamic Games **5**, pp. 387–404.

[Petrosjan & Zaccour 2003] L. Petrosjan, G. Zaccour (2003). *Time-consistent Shapley value allocation of pollution cost reduction*. Journal of Economic Dynamics and Control **27**, no. 3, pp. 381–398.

[Petrosjan & Zakharov] L. Petrosjan, V. Zakharov (1997). *Mathematical models in environmental policy analysis. Nova Science Publishers*. New York, USA, 262 pp.

[Ray & Vohra 1997] D. Ray, R. Vohra (1997). *Equilibrium binding agreements*. Journal of Economic Theory 73, pp. 30–78.

[Ray &Vohra 1999] D. Ray, R. Vohra (1999). *A Theory of Endogeneous Coalition Structures*. Game and Economic Behavior **26**, pp. 286–336.

[Rubio & Casino] S. Rubio, B. Casino (2001). *International Cooperation in Pollution Control*. Working paper 2001-21, Valenciano de Investigaciones Económicas.

[Rubio & Ulph 2002] S. Rubio, A. Ulph (2002). *A simple dynamic model of international environmental agreements with a stock pollutant*. Discussion Paper in Economics and Econometrics 0209, Department of Economics, University of Southampton.

[Rubio & Ulph 2007] S. Rubio, A. Ulph (2007). *An infinite-horizon model of dynamic membership of international environmental agreements*. Journal of Environmental Economics and Management **54**, issue 3, pp. 296–310.

[Strotz 1955] R. H. Strotz (1955–1956). *Myopia and Inconsistency in Dynamic Utility Maximization*. The Review of Economic Studies **23**, no. 3, pp. 165 –180.

[Weikard 2005] H.-P. Weikard (2005). *Cartel Stability Under an Optimal Sharing Rule*. FEEM Working Paper No.77.05. Venice: FEEM.

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

[Weikerd & Dellink 2008] H.-P. Weikard, R. Dellink (2008). *Sticks and Carrots for the Design of International Climate Agreements with Renegotiations* FEEM Working Paper No. 26.2008. Venice: FEEM.

[Wirl 2004] F. Wirl (2004). *International greenhouse gas emissions when global warming is a stochastic process*. Applied Stochastic Models in Business and Industry **20**, no. 2, pp. 95–114.

[Yi 1997] S.-S. Yi (1997). *Stable Coalition Structures with Externalities*. Games and Economic Behavior **20**, pp. 201–237.

[Yi 2003] S.-S. Yi (2003). *Endogeneous Formation of Economic Coalitions: A Survey of the Partition Function Approach*. (Eds. C. Carraro) The Endogeneous Formation of Economic Coalitions, Edvard Elgar, Cheltenham, UK.

[Zakharov 1988] V. Zakharov (1988). *Regularization and time-consistency of dynamic solutions in hierarchical differential games*. (Russian) Vestnik Leningrad. Universiteta, ser. 1, vyp. 2, no. 8, pp. 45–53.

[de Zeeuw 2008] A. de Zeeuw (2008). *Dynamic Effects on the Stability of International Environmental Agreements*. Journal of Environmental Economics and Management **55**, no. 2, pp. 163–174.

YHTEENVETO (FINNISH SUMMARY)

Ympäristöongelmien kehittyessä alati uhkaavammiksi ovat myös poliittiset lähestymistavat tilanteen korjaamiseksi muuttuneet alati kansainvälisemmiksi. Viimeisen vuosikymmenen kuluessa joukko monenkeskisiä sopimuksia ympäristön suojelemiseksi on pantu alulle, vaikkakin niiden tehokkuus ja olemassaolon jatkuvuus on suuressa vaarassa. Useat edellä mainittuihin sopimuksiin kohdistuneet haasteet ovat selitettävissä positiivisen ulkopuolisuuden käsitteellä, joka tarkoittaa että tietyt maat hyötyvät naapurimaidensa ympäristöystävällisistä toimista ponnistelematta itse asian hyväksi. Näin ollen, jos sopimus asettaa oman edun kannalta ristiriitaisia rajoituksia maiden teknologiselle ja ekonomiselle kehittymiselle, johtaa tämä eräänlaiseen vapaamatkustukseen osallistumisen ja sääntöjen noudattamisen suhteen. Monenkeskisen yhteistyön saavuttamiseksi osapuolten strategiset intressit täytyy tuntea. Tässä työssä peliteoria on valittu näiden intressien, ja sen kuinka ne vaikuttavat kansainvälisten ympäristösopimusten (international environmental agreement, IEA) muodostamiseen, suunnitteluun ja toimivuuteen, analyysin välineeksi. Luotuamme katsauksen peliteoreettisiin menetelmiin IEA-kentässä, tarkastelemme tiettyjä IEA:iin liittyviä näkökulmia, kuten päästösupistusten tavoitteita, jäsenyyden asemaa (allekirjoittajavaltio vs. vapaamatkustaja), sekä erilaisia mekanismeja, joilla maita voitaisiin motivoida ottamaan osaa sopimuksiin. Tarkastelemme erityisesti seuraavaa kysymystä: jos osanottajilla ja vapaamatkustajilla tietyssä IEA sopimuksessa on kannustimena muuttaa statustaan sopimuksen voimassaoloaikana, muodostaako tämä uhan sopimuksen noudattamiselle? Kehitämme metodologian IEA:n vakauden saavuttamiseksi, ja osoitamme että tätä metodologiaa hyödyntäen IEA:n sopijapuolet noudattavat päästösupistusten tavoitteita. Lisäksi, arvioimme sopimusten vaikutuksia sopijamaiden hyvinvointiin, sekä globaalin tason saastumiseen.