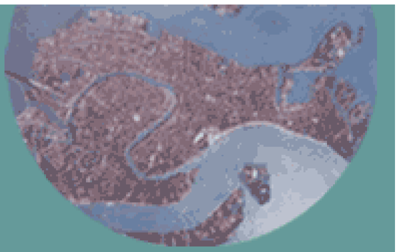
EAERE FEEM VIU
European Summer School
In Resources and Environmental Economics
POLITICAL ECONOMY OF THE ENVIRONMENT
September, 1st - 6th, 2003



The effect of membership rules and voting schemes for the success of international climate agreements

Authors:

J. Carlos Altamirano-Cabrera
Ekko Van Ierland
Wageningen University
The Netherlands

Michael Finus
Hagen University
Germany

The Effect of Membership Rules and Voting Schemes for the Success of International Climate Agreements[#]

Michael Finus [*], Juan-Carlos Altamirano-Cabrera [+] and Ekko van Ierland [+]

* Hagen University
Department of Economics
Institute of Economic Theory
Profilstr. 8
58084 Hagen
Germany

+ Wageningen University
Department of Social Sciences
Environmental Economics and Natural Resources Group
Hollandseweg 1
6706 KN Wageningen
The Netherlands

Abstract

We empirically test the role of membership rules and voting schemes for climate change coalitions with the STAbility of COalitions model (STACO). The model comprises 12 world regions and captures important dynamic aspects of the climate change problem. We apply three stability concepts that capture the notion of open membership and exclusive membership with majority and unanimity voting. We show that exclusive membership leads to superior outcomes than open membership and that unanimity voting is preferable to majority voting in welfare and ecological terms. Our results suggest restricting membership in future international environmental agreements and they provide a rationale for unanimity voting as applied in many international organizations.

JEL-Classification: C72, D72, H41, Q25

Keywords: international environmental agreements, Kyoto-Protocol, public choice, non-cooperative game theory

---

[#] The authors acknowledge the support of Rob Dellink, Eligius Hendrix, Niels Olieman, and Hans-Peter Weikard in the analysis.

1. Introduction

International environmental agreements (IEAs) are examples of collective action to tackle global problems such as global warming. Game theoretic analyses of the formation of IEAs stress the difficulties in designing self-enforcing treaties because of free-riding. The presence of a strong free-rider incentive prevents most IEAs of being stable and/or effective. The Kyoto Protocol is a clear example of this problem. For studying the problems of forming large and effective coalitions non-cooperative game theory has proved to be a very fruitful approach (e.g., Barrett 1994 and 1997, Bauer 1992, Carraro/Siniscalco 1993, Hoel 1992, Hoel/Schneider 1997 Jeppesen/Andersen 1998 and Rubio/Ulph 2001).[1] Key results that emerge from this literature are: a) only small coalitions are stable and b) whenever full cooperation (global optimum) would generate large global welfare gains compared to a no cooperation (Nash equilibrium), stable partial cooperation achieves only little.

A stability concept that has been widely used in non-cooperative game theory is "internal and external stability". Internal stability means that no coalition member has an incentive to leave the agreement to become a non-signatory. External stability means that no non-signatory has an incentive to join the agreement. This stability definition implies that countries can freely choose to join or leave the agreement. In particular, the definition of external stability means that there is no restriction on membership. Thus, coalition formation may be seen as an open membership game.

Open membership seems in line with evidence on IEAs. Almost all protocols of major IEAs have no provision that restricts membership. Moreover, intuition and results of the public goods literature suggest that global welfare increases with participation in an agreement and therefore any restriction on membership would hamper the effectiveness of IEAs. However, those results have been derived without considering the restriction that IEAs have to be self-enforcing. Hence, in the presence of free-rider incentives, it seems worthwhile to study the effect on coalition formation when membership is restricted. For instance from the empirical study of Botteon/Carraro (1997) on global warming it appears that some coalitions are internally but not externally stable. This suggests that some of these coalitions could be stabilized if coalition members had the opportunity to deny accession of new potential entrants. Also recent theoretical results by Finus/Rundshagen (2003) obtained from a general framework

---

[1] There is also a large literature applying cooperative game theory to study stability of IEAs (see, e.g., Chander/Tulkens 1995, 1997 and Germain et al. 2000 and an overview in Finus 2001 and 2003a). This literature is mainly normatively oriented and has focused on measures to stabilise the efficient grand coalition implementing a socially optimal emission or abatement vector.

suggest that exclusive membership may help to stabilize IEAs. The reason is that - though still internal instability poses a problem to IEAs - external instability is less of a problem because members of an IEA under exclusive membership can better control the accession of non-signatories that may upset a coalition equilibrium. However, Finus/Rundshagen (2003) point out that at a general level it is not possible to conclude what “more stability” means in terms of the success of IEAs. Therefore, it is the purpose of this paper to study by means of an empirical model the implications of exclusive membership in terms of stability, global welfare and ecological variables (global emissions and stock of greenhouse gases). We consider two exclusive membership rules – one with (simple) majority voting and a second with unanimity voting.

Hence, our analysis is in the tradition of the public choice theory that analyzes the effect of different voting procedures for instance in parliament and various national and international organizations on the provision of public goods.[2] By and large, this literature has focused on voting schemes where a given number of participants decide either on various policy options (platforms) or on the level of a policy instrument. In this context, voting means for instance deciding on new trade laws or whether and by how much tariff barriers within WTO should be reduced, passing a resolution within the United Nation Council and deciding on the implementation of antitrust laws within the European Union. Typical voting procedures are simple majority voting and qualitative majority voting as has emerged in recent years in the European Union, as well as unanimity voting as for instance in the United Nations and the WTO.

In contrast, our model makes a particular assumption on the policy level and implicitly on the policy instrument, but analyzes voting on membership. As will be outlined in more detail in section 2, we assume that coalition members chose their abatement strategies such as to maximize the aggregate payoff to their coalition and that there are no compensation payments. This implies that abatement strategies are not only cost-efficient but also optimal from cost-benefit considerations of the coalition. Though we do not pay special attention to policy instruments, our assumption de facto implies that an efficient policy instrument (e.g., a uniform tax) is implemented within a coalition without compensation payments or any kind of redistribution of the gains from cooperation. Thus, in our model inefficiency stems from free-

---

[2] In the general context see for instance Buchanan/Tullock (1962), Mueller (2003) and McNutt (1996) and in the context of environmental policy Böhringer/Vogt (2003), Dijkstra (1999), Hahn (1989), Hillman/Ursprung (1994), Michaelowa (1998), Michaelowa/Greiner (1996), Schneider/Volkert (1999), Sandler (1992) and Yandle (1999).

riding. As mentioned above, we consider stability under two membership rules (open versus exclusive membership) and for exclusive membership we consider two voting schemes: majority voting and unanimity voting. Our voting schemes under exclusive membership are frequently applied in international treaties (that are not necessarily public good agreements), like the NATO, the European Union, WTO or the United Nations. For instance, NATO and WTO vote by unanimity on new members. Also a new permanent member in the Security Council is only accepted if there is no veto. The decision on accession of ten new member states in the European Union is an example of different voting rules. First, accession had to be accepted by the European Parliament by a simple majority and subsequently the European Council had to approve the accession by unanimity (Euractiv 2003).

We believe that our paper extends previous work in several directions. First, we consider stability of IEAs not only under open membership but also under exclusive membership and different voting rules. Second, we confirm several conclusions of theory that have been derived under very restrictive assumptions (e.g., symmetric countries or heterogeneous countries with two types of countries). Third, our analysis captures important dynamic aspects of the climate change problem that have been ignored by the non-cooperative game theoretic literature on coalition formation by and large.[3] Fourth, in contrast to other empirical studies (e.g., Botteon/Carraro 1997), our analysis captures a sufficient number of different actors (12 regions) that makes the strategic interaction in the context of global warming interesting.

In what follows, we lay out the game theoretical part of the model in section 2 and the empirical part in section 3. In section 4, we discuss ecological and welfare aspects of coalition formation of our base case scenario and in section 5 we report on results of various sensitivity analyses. Section 6 summarizes the main findings, draws policy conclusions and concludes with some remarks about future research issues.

2. Theoretical Background of the Model

Coalition formation is modeled as a *two-stage game*. In the *first stage* countries or regions decide on their *membership* in a coalition; in the *second stage* coalition members choose their *abatement strategies*. In the *first stage* we assume that countries can choose between two membership strategies: strategy $\sigma_i = 0$ means "I do not want to sign the agreement" and $\sigma_i = 1$ means "I want to become a member of a climate treaty". Technically, this implies that countries that announce $\sigma_i = 0$ form a singleton coalition and those that announce $\sigma_i = 1$

---

[3] Exceptions are for instance Rubio/Ulph (2001) and Tol (2001).

become members of a non-trivial coalition (i.e., a coalition of at least two members). More formally, we have:

Definition 1: Stage 1 of the Coalition Formation Game

Let i denote a particular country, $i \in I = \{1, \ldots, N\}$*, and let a particular membership strategy of country i be the message* σ_i *and its strategy set be given by* $\Sigma_i = \{0, 1\}$, $\Sigma = \Sigma_1 \times \Sigma_2 \times \ldots \times \Sigma_N$*, and denote* c^i *the coalition to which i finally belongs, then*

$$c^i = \begin{cases} \{i\} \text{ if } \sigma_i = 0 \\ \{j / \sigma_j = 1\} \text{ if } \sigma_i = 1 \end{cases}.$$

A coalition structure $c = (c^1, \ldots, c^M)$ *is a partition of countries where a particular coalition is denoted by* c^k, $k \in \{1, \ldots, M\}$, $c^k \cap c^l = \varnothing \; \forall k \neq l$, $\bigcup c^i = I$ *and* $c \in C$ *where C is the set of coalition structures.*

Due to the restriction to two strategies, notation can be simplified and we may write $c = (c^S, 1, \ldots, 1)$ instead of $c = (c^1, \ldots, c^M)$. Moreover, when country i and j are members of c^S in coalition structure $c = (c^S, 1, \ldots, 1)$, we will call this simply a coalition between country i and j; and if $c^S = \{i\}$ we call this the *"singleton coalition structure"* and if $c^S = I$ the *"grand coalition"*.

In the *second stage* countries choose their abatement strategies based on the following payoff function:

$$[1] \qquad \pi_i(q) = \sum_{t=1}^{T} (1 + r_i)^{-t} (B_{it}(q_t) - AC_{it}(q_{it}))$$

where T denotes the time horizon, t=1, ..., T, r_i is the discount rate of country i, B_{it} are benefits from global abatement $q_t = \sum_{i=1}^{N} q_{it}$, AC_{it} are abatement costs from individual abatement q_{it} and q is an abatement vector of dimension $N \times T$. Benefits from global abatement are derived from reduced environmental damages caused by greenhouse gas emissions. We make the standard assumption: $\forall$ $i \in I$, $q_{it} \in [0, e_{it}^{BAU}]$ and at each time t: $B_{it}^{'} > 0$, $B_{it}^{''} \leq 0$, $AC_{it}^{''} > 0$ and $AC_{it}^{''} > 0$ where primes denote first and second derivatives and e_{it}^{BAU} is the emission level in the business-as-usual scenario with no abatement.

In section 3 we will lay out in detail how global abatement relates to global emissions and the stock of greenhouse gases and how this affects payoffs. At this stage it suffices to note that we follow the standard assumption in coalition theory and presume that countries belonging to the same coalition maximize the aggregate payoff to their coalition (Bloch 1997). The equilibrium abatement strategy vector q^* for coalition structure c is derived as a Nash equilibrium between coalitions. In our context this implies that non-signatories maximize their own payoff and signatories maximize the sum of payoffs of the members of the agreement. Hence, following the terminology of Chander/Tulkens (1995 and 1997) q^* may also be called partial Nash equilibrium between the members of the agreement and the remaining countries. More formally, we have:

Definition 2: Stage 2 of the Coalition Formation Game

Fix a coalition structure $c=(c^S,1,...,1)$ *let* $\upsilon_i(c)=\pi_i(q^*)$ *and assume that signatories* $i\in I^S$ *jointly maximize the aggregate payoff to their coalition* c^S *and each non-signatory* $j\in I^{NS}$ *maximizes his own payoff* ($I^{NS}\cap I^S=\varnothing, I^{NS}\cup I^S=I$)*. Let* q^S *denote the abatement strategy vector of signatories and* q_j^{NS} *the abatement strategy vector of non-signatory j,* $q_i\in Q_i$, $Q=Q_1\times...\times Q_N$*, and assume that the equilibrium abatement vector* $q^*=(q^{S*},q^{NS*})=(q^{S*},q^{I\backslash I^S*})=(q_j^*,q^{I\backslash j*})$ *satisfies:*

$\forall\, c^S\in C,\ \forall\, q^S\in Q^S:\ \sum_{i\in c^S}\pi_i(q^{S*},q^{I\backslash I^S*})\geq\sum_{i\in c^S}\pi_i(q^S,q^{I\backslash I^S*})$ *and*

$\forall j\in I^{NS},\ \forall\, q_j\in Q_j:\ \pi_j(q_j^*,q^{I\backslash j*})\geq\pi_j(q_j,q^{I\backslash j*})$ *where* q^* *is assumed to be a unique interior equilibrium.*

Definition 2 implies that the payoff of country i, $\upsilon_i(c)$, can solely be identified by a coalition structure c. Signatories behave cooperatively within their coalition but non-cooperatively against non-signatories. Hence, abatement strategies within coalition c^S are efficiently chosen. Consequently, the singleton coalition structure (grand coalition) implies an equilibrium abatement strategy vector corresponding to the "classical" Nash equilibrium (global optimum). Thus, the highest global payoff will be obtained in the grand coalition, the lowest in the singleton coalition structure and payoffs in between in any other coalition including more than one and less than N members. For the calibration of the payoff functions (also called net benefit functions below)- on which we report in section 3 - it turns out that q^* (for all coalition structures $c\in C$) is unique and lies well within the boundaries of the abatement space as defined above ($q_{it}\in[0,e_{it}^{BAU}]$).

From definition 2 it is evident that because the strategy in the second stage is fixed, the entire coalition formation game reduces to one single stage. This reduced stage game looks as follows: each country chooses its membership strategy, σ_i, the strategy vector $\sigma=(\sigma_1,...,\sigma_N)$ is mapped into coalition structure c, leading to payoff vector $\upsilon=(\upsilon_1,....,\upsilon_N)$. Hence, we can define stability in terms of $\upsilon_i(c(\sigma))$. For brevity, however, we write only $\upsilon_i(\sigma)$ in the following if no misunderstanding is possible.

The first definition of stability is the standard definition of internally and externally stable coalition structures (I&E-CS) applied to our context.

Definition 3: Stability under Open Membership

Denote the strategy of country i by σ_i *where coalition structure* c^* *is generated by strategy vector* σ^**, then* c^* *is called stable if* $\forall i \ \upsilon_i(\sigma_i^*, \sigma_{-i}^*) \geq \upsilon_i(\sigma_i, \sigma_{-i}^*) \ \ \forall \sigma_i \neq \sigma_i^*$.

Definition 3 implies that a country that announces $\sigma_i^* = 1$ (signatory) will not leave the agreement to become non-signatory by announcing $\sigma_i = 0$ if this implies a lower valuation. This is the classical condition of internal stability. By the same token, a non-signatory that announces $\sigma_j^* = 0$ will not join the agreement to become a signatory by announcing $\sigma_j = 1$ if this implies a lower payoff. This is the classical definition of external stability. It implies that a non-signatory can freely join the agreement without the consent of signatories. Therefore, we call this type of stability "stability under open membership".

From Definition 3 it is evident why we - different from the main stream of the literature - model the first stage of the coalition formation game as an announcement game. In our setting, an equilibrium always exists because the singleton coalition structure is stable by definition. The reason is simple: suppose each country announces $\sigma_i = 0$, then no single country can induce another coalition structure by changing its announcement. In contrast, applying the standard definition of I&E-CS to determine stable coalition structures, the singleton coalition structure may not be externally stable and if no other coalition structure satisfies the conditions of internal and external stability an equilibrium may fail to exist.

The next two definitions capture the notion of exclusive membership. Definition 4 assumes that an agreement is not only stable if no non-signatory wants to join the agreement but also if a majority of signatories is against accession of a new entrant. Definition 5 imposes even stricter rules for accession: a new potential entrant (non-signatory) can accede to an agreement *if and only if* there is unanimous consent by all signatories about its accession.

Definition 4: Stability under Exclusive Membership and Majority Voting

Denote the set of countries announcing $\sigma_i = 0$ *by* I^{NS} *(non-signatories), the set of countries with m members announcing* $\sigma_i = 1$ *by* I^S *(signatories),* $I^{NS} \cap I^S = \emptyset$*,* $I^{NS} \cup I^S = I$*), and let* I^{SC} *be a subset of n players belonging to coalition* I^S*, then* c^* *generated by* σ^* *is called stable if* $\forall \ i \in I$ *and* $\sigma_i \neq \sigma_i^*$ *:* $\upsilon_i(\sigma_i^*, \sigma_{-i}^*) \geq \upsilon_i(\sigma_i, \sigma_{-i}^*)$ *or* $i \in I^{NS}$ *and* $\sigma_i \neq \sigma_i^* = 0$ *:* $\upsilon_i(\sigma_i^*, \sigma_{-i}^*) < \upsilon_i(\sigma_i, \sigma_{-i}^*)$ *and* $\exists \ \ I^{SC} \subset I^S$ *of size n>m/2, with strategy* $\sigma_j^* = \sigma_i = 1$*:* $\upsilon_j(\sigma_i^*, \sigma_{-i}^*) > \upsilon_j(\sigma_i, \sigma_{-i}^*)$.

Definition 5: Stability under Exclusive Membership and Unanimity Voting

Denote the set of countries announcing $\sigma_i = 0$ *by* I^{NS} *and the set of countries announcing* $\sigma_i = 1$ *by* I^S*, then* c^* *generated by* σ^* *is called stable if* $\forall\ i \in I$ *and* $\sigma_i \neq \sigma_i^*$*:* $\upsilon_i(\sigma_i^*, \sigma_{-i}^*) \geq \upsilon_i(\sigma_i, \sigma_{-i}^*)$ *or* $i \in I^{NS}$ *and* $\sigma_i \neq \sigma_i^* = 0$*:* $\upsilon_i(\sigma_i^*, \sigma_{-i}^*) < \upsilon_i(\sigma_i, \sigma_{-i}^*)$ *and* $\exists\ j \in I^S$ *with strategy* $\sigma_j^* = \sigma_i = 1$*:* $\upsilon_j(\sigma_i^*, \sigma_{-i}^*) > \upsilon_j(\sigma_i, \sigma_{-i}^*)$*.*

When comparing the three definitions (Definitions 3 to 5) it is evident that all definitions require "internal stability" of a coalition structure but differ in the definition of "external stability". From Definition 3 to Definition 5 the “degree of exclusivity” rises from open membership to exclusive membership and is higher under unanimity than under majority voting. Hence, it is more difficult to upset an equilibrium coalition structure under exclusive membership than under open membership and under exclusive membership it is more difficult to upset an equilibrium under unanimity voting than under majority voting. Thus, if we abbreviate the set of equilibrium coalition structures under open membership, exclusive membership and majority voting and exclusive membership and unanimity voting by C(OM), C(EM-M) and C(EM-U), then C(OM)$\subset$C(EM-M)$\subset$C(EM-U) must hold by simple theoretical reasoning. Therefore, the interesting question is: what does "more or less stability" imply for coalition formation and in particular what are the ecological and welfare implications? At a theoretical level this question can only be answered for very restrictive assumptions such as symmetric payoff functions (see, e.g., Finus (2002) and Finus/Rundshagen 2001). Therefore, it is important to analyze this question in an empirical context as we do in sections 4 and 5.

3. Empirical Background of the Model

3.1 Introduction

In this section we describe the calibration of payoff function [1]. The philosophy behind the construction of our empirical model comprises two items. First, the model must be simple enough to be tractable for a game theoretical analysis. Nevertheless, the model should reflect important results and features of climate models in terms of the development of global emissions and concentration over some time period. Therefore, we base our calibration in this respect on the widely known DICE-model by Nordhaus (1994). Second, in order to make the model interesting for a game theoretic analysis, there should be a sufficient number of different players. We consider 12 world regions. Since this requires disaggregated information on benefit and abatement cost functions we rely on damage cost estimates of Fankhauser (1995) and Tol (1997) and abatement cost estimates of Ellerman/Decaux (1998). We set up an empirical model that we call *sta*bility of *co*alitions model, henceforth abbreviated STACO.

STACO captures important dynamic aspects of climate change but is de facto a finitely repeated game with stationary abatement strategies.

In the following we proceed in five steps. First, we describe the relation between emissions and stock of greenhouse gases. Second, we discuss damages implied by concentration. Third, we show how we derive benefit functions from damage cost functions. Fourth, we report about the calibration of the abatement cost functions. Fifth, we discuss the implications of the first four steps for our payoff function and computations of valuations for different coalition structures. All parameters are reported in the Appendix; a detailed description of the model is available from the authors upon request (Dellink et al. 2003).

3.2 Emissions and Stock of Greenhouse Gases

In our analysis we focus on carbon dioxide, but the exogenous level of other greenhouse gases is included in the calibration of the damage cost function (Nordhaus 1994). For the development of emissions and the stock of carbon dioxide in the business-as-usual-scenario (BAU) we base our calibration on the market scenario in DICE. This scenario assumes no emission reduction, though there is a feedback between the environment and the economy. In DICE, global emissions grow non-constantly over time. However, it turns out that a linear specification of uncontrolled global emissions (e_t) provides a good fit for the development of the stock of carbon dioxide:

$$[2] \quad e_{t+1} = e_t + d_E$$

where d_E denotes the uncontrolled annual increment that is added to global emissions, $e_t = \sum_{i=1}^{N} e_{it}$. Our analysis starts in 2010 and covers a time period of 100 years in order to capture the long-run effects of the global warming problem. Thus, with reference to equation [1], t=2011, ..., 2110. For emissions in 2010 we choose the value of DICE, which amounts to 11.96 gigatons CO_2. We estimate [2] using OLS-regression. This gives $d_E = 0.153$.

The stock of carbon dioxide in the atmosphere at time t is expressed in the standard way by the following equation:

$$[3] \quad M_t\left(q_{2011},\ldots,q_t\right) = M_{pre-ind} + \left(1-\delta\right)^{(t-2010)} \cdot \left(M_{2010} - M_{pre-ind}\right) + \sum_{s=2011}^{t} \left(\left(1-\delta\right)^{t-s} \cdot \omega \cdot \left(e_s - q_s\right)\right)$$

That is, the stock at time t, M_t, depends on global abatement from time t=2011 onwards, q_{2011}, ..., q_t where $q_t = \sum_{i=1}^{N} q_{it}$. More specifically, the stock depends on three terms. The first term is the pre-industrial stock, $M_{pre-ind}$, which is 590 gigatons CO_2 according to DICE. This stock remains constant over time and may be interpreted as the "natural equilibrium".

The second term is the stock in 2010 in excess of the pre-industrial stock that decays with a rate δ per annum. The "natural removal or decay rate" as well as the stock in 2010 are taken from DICE and are $\delta = 0.00866$ and $M_{2010} = 835$ gigatons CO_2, respectively. The third term constitutes that part of the stock that is due to global (BAU-) emissions e_s, which grow according to [2], minus global abatement after 2010, q_s. The airborne fraction of total net emissions (BAU-emissions minus abatement) that remains in the atmosphere is 64 percent ($\omega = 0.64$) according to DICE, which decays with rate $\delta = 0.00866$ per annum. In the BAU-scenario with no abatement, the uncontrolled stock according to [3] in 2110 is 1585 gigatons whereas the corresponding value taken from DICE is 1576 gigatons. This stresses that our approximation in [2] works well.

If we denote the uncontrolled stock at time t by $M_t(0)$, then [3] can be rewritten:

$$[4] \qquad M_t\left(q_{2011},\ldots,q_t\right) = M_t(0) - \sum_{s=2011}^{t} (1-\delta)^{t-s} \cdot \omega \cdot q_s \,.$$

which simplifies if we assume q_{it} (and hence also q_t) constant over time. For the stock of CO_2 in 2110 this leads to:

$$[5] \qquad M_{2110}(q) = M_{2110}(0) - \left[\sum_{t=2011}^{2110} (1-\delta)^{2110-t} \cdot \omega\right] \cdot \frac{q}{100}$$

where $q = \sum_{t=2011}^{2110} q_t$, the term in brackets is a constant equal to 42.9 and $M_{2110}(0)$=1585 gigatons CO_2 as reported above.

3.3 Global Damage Cost Function

In DICE global damages depend on world temperature increase, ΔT_t, global GDP, Y_t, and parameter γ_D that measures the impact on GDP due to an increase in temperature of 3 degrees Celsius compared to the pre-industrial level.

$$[6] \qquad D_t = \gamma_D \cdot \left[\frac{\Delta T_t}{3}\right]^2 \cdot Y_t$$

However, in order to establish a direct link between concentration and damages, we follow Germain and Van Steenberghe (2001), who use the following approximation of the full climate module:

$$[7] \qquad \Delta T_t = \eta \cdot \ln\left(\frac{M_t}{M_{pre-ind}}\right)$$

where η is a parameter. Substituting [7] into [6], gives:

$$[8] \qquad D_t = \left(\frac{\gamma_D}{9}\right) \cdot \left[\eta \cdot \ln\left(\frac{M_t}{M_{pre-ind}}\right)\right]^2 \cdot Y_t$$

In DICE it is assumed that a doubling of the carbon dioxide concentration ($2 \cdot M_{pre-ind}$) leads to an increase in temperature of 3 degrees.[4] Thus from [7], $\eta = 3/\ln(2)$, and γ_D can be interpreted as damages in percentages of GDP for a doubling of concentration:

$$[9] \qquad D_t = \left[\frac{1}{\ln(2)} \cdot \ln\left(\frac{M_t}{M_{pre-ind}}\right)\right]^2 \cdot (\gamma_D \cdot Y_t)$$

Though this damage function is non-linear, it can be approximated by a linear function in the relevant range of our study, that is, between the stock in 2010 (1.4 times pre-industrial level) and the estimated uncontrolled level in 2110 (3.5 times pre-industrial level):

$$[10] \qquad D_t = \left[\gamma_1 + \gamma_2 \cdot \left(\frac{M_t}{M_{pre-ind}}\right)\right] \cdot (\gamma_D \cdot Y_t)$$

where γ_1 and γ_2 are calculated via OLS-regression. Further manipulation that considers the fact that (i) a doubling of concentration occurs between 2055 and 2065 in DICE and also in our approximation, (ii) the undiscounted GDP in this period is 70284 billion US$ and (iii) $M_{pre-ind} = 590$ gigatons CO_2, we derive:[5]

$$[11] \qquad D_t = \gamma_D \cdot (\varphi_1 + \varphi_2 \cdot M_t)$$

where $\varphi_1 = \gamma_1 \cdot Y_{2061} = -140146$ billion US$ and $\varphi_2 = \gamma_2 \cdot (1/M_{pre-ind}) \cdot Y_{2061} = 178.331$ billion US$ per Gton.

3.4 Derivation of Global and Regional Benefit Functions

Since we prefer to compute payoffs in terms of net benefits and not in terms of total costs, we express benefits in the form of reduced damages due to abatement. Due to the assumption of stationary abatement strategies, we can express benefits in year t as a function of total abatement over the entire period, $q = \sum_{t=2011}^{2110} q_t$. Noting that [11] reads $D_t(M_t(q)) = \gamma_D \cdot (\varphi_1 + \varphi_2 \cdot M_t(q))$ if abatement is explicitly accounted for, we derive for benefits from global abatement in year t, $B_t(q)$:

---

[4] This is based on an exogenous additional impact of other greenhouse gases on radiative forcing (see Nordhaus 1994).

[5] All market values are expressed in billion US$ of 1985 using the deflator provided by NASA (2002). This applies to damages, benefits and abatement costs.

$$
\begin{aligned}
B_t(q) &= D_t\left(M_t(0)\right) - D_t\left(M_t(q)\right) \\
\text{[12]} \qquad &= \gamma_D \cdot \left[\varphi_1 + \varphi_2 \cdot M_t(0)\right] - \gamma_D \cdot \left[\varphi_1 + \varphi_2 \cdot M_t(q)\right] \\
&= \gamma_D \cdot \varphi_2 \cdot \left(M_t(0) - M_t(q)\right)
\end{aligned}
$$

which indicates that the intercept φ_1 has no effect on the benefit function. Summing over all periods, discounting benefits with a discount rate of 2 percent, inserting $\varphi_2 = 178.331$ from above gives total benefits $TB(q) = \gamma_D \cdot 1385.1 \cdot q$ and marginal total benefits $MTB(q) = \gamma_D \cdot 1385$. Nordhaus (1994) assumes for the scale parameter γ_D a value of 0.0133, that is, damages amount to 1.33 percent of GDP. However, it is known that the DICE value is relatively low. Therefore, we use the more recent estimate of Tol (1997) who estimates damage costs of 2.7 percent of GDP for a doubling of concentration and hence $\gamma_D = 0.027$. This leads to $TB(q) = 37.40 \cdot q$, implying discounted marginal global benefits of 37.40 US$ per ton CO_2 ($MTB(q) = 37.4$). This figure is in line with results by Plambeck and Hope (1996) who report that their best estimates of marginal global benefits in a regional scenario fall within the range of 10 to 48 US$ per ton CO_2.

In a final step, we have to allocate global benefits from reduced environmental damages to the various world regions based on the assumption that: $TB_i(q) = s_i \cdot TB(q)$ (and hence $MTB_i(q) = s_i \cdot MTB(q)$) where s_i is the share of region i. We consider 12 regions: USA (USA), Japan (JPN), European Union (EEC), other OECD countries (OOE), Eastern European countries (EET), former Soviet Union (FSU), energy exporting countries (EEX), China (CHN), India (IND), dynamic Asian economies (DAE), Brazil (BRA) and "rest of the world" (ROW).[6] The allocation is a difficult task since no source of damage cost estimates is available that exactly matches with our regions. However, Fankhauser (1995) and Tol (1997) are two sources that come relatively close to our regional specification. Hence, we adjust their

---

[6] EEC comprises the 15 countries of the European Union as of 1995. Other OECD countries (OOE) includes among other countries Canada, Australia and New Zealand. Eastern European countries (EET) include for instance Hungary, Poland, and Czech Republic. Energy Exporting Countries (EEX) includes for instance the Middle East Countries, Mexico, Venezuela and Indonesia. Dynamic Asian economies (DAE) comprises South Korea, Philippines, Thailand and Singapore. Rest of the World (ROW) includes for instance South Africa, Morocco and many countries in Latin America and Asia. For details see Babiker et al. (2001).

estimates for our purposes.[7] This gives rise to the shares displayed in the third column in Table 1.

Table 1: Benefit and Abatement Cost Parameters

Regions	Emissions in 2010 (Gton)	Share of global benefits s_i	Abatement cost parameter β_i	Abatement cost parameter α_i
1 USA	2.42	0.226	0.0005	0.00398
2 JPN	0.56	0.173	0.0155	0.18160
3 EU	1.4	0.236	0.0024	0.01503
4 OOE	0.62	0.035	0.0083	0
5 EET	0.51	0.013	0.0079	0.00486
6 FSU	1	0.068	0.0023	0.00042
7 EEX	1.22	0.030	0.0032	0.03029
8 CHN	2.36	0.062	0.00007	0.00239
9 IND	0.63	0.050	0.0015	0.00787
10 DAE	0.41	0.025	0.0047	0.03774
11 BRA	0.13	0.015	0.5612	0.84974
12 ROW	0.7	0.068	0.0021	0.00805
WORLD	11.96	$\sum s_i = 1$		

From Table 1 it is evident that shares of OECD countries are higher than those of IND, DAE, BRA and ROW.

3.5 Derivation of Abatement Cost Functions

For the specification of the abatement cost function we rely on estimates of the EPPA model that are reported in Ellerman and Decaux (1998). They assume an annual abatement cost function of the following form:

$$[13] \quad AC_{it}(q_{it}) = \tfrac{1}{3} \cdot \alpha_i \cdot (q_{it})^3 + \tfrac{1}{2} \cdot \beta_i \cdot (q_{it})^2$$

We can use their estimates but have to adjust their figures in four respects. First, we have to account for the fact that their abatement cost estimates are in million US$ per megaton greenhouse gas reduction whereas our unit of measurement is billion US$ per gigaton. Second, we replace q_{it} by $q_i/100$ in [13] because we assume stationary strategies ($q_{i,2011} = \ldots = q_{i,2110}$). Third, they estimate a negative value for the parameter α_i for OOE. Since this would cause problems for computations, we set α_i=0 in this case and re-estimate β_i for OOE. All estimates are displayed in the last two columns in Table 1. Fourth, in our model abatement means

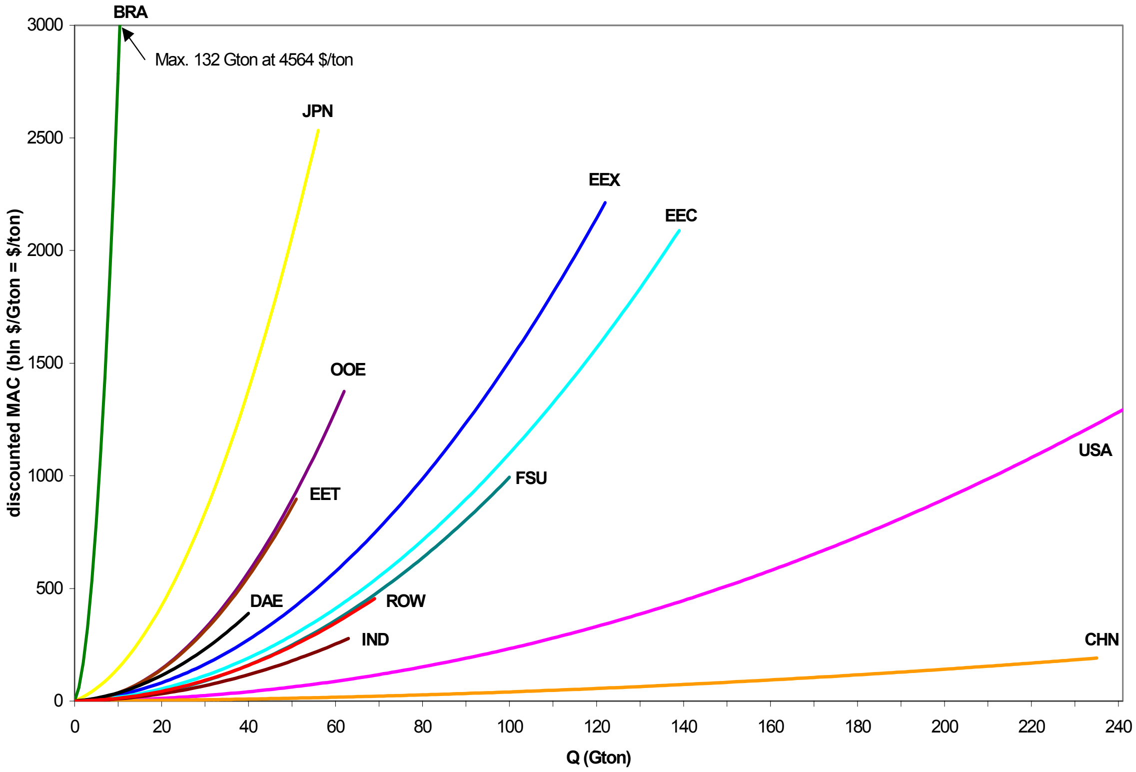
---

[7] Because of space limitations the interested reader is referred to our empirical background paper that is available upon request from the authors. There we lay out in detail how we derive shares s_i from Fankhauser (1995) and Tol (1997).

emission reduction with respect to BAU-emissions. Thus, we allocate total initial emissions of 11.96 gigatons (see section 3.2) to the 12 regions, using the shares of Ellerman and Decaux (1998). This gives the numbers in the second column in Table 1. This implies that we assume not only global emissions to grow linearly with d_E (see equation [2]) but also regional emissions, however, according to their shares in global emissions.

In order to derive total abatement costs of region i, $TAC_i(q_i)$, we sum [13] over t=2011, ..., 2110 and discount with discount rate r, $TAC_i(q_i) = \sum_{t=2011}^{2110}(1+r)^{-(t-2010)} AC_{it}(q_i)$. This implies that we assume the same abatement cost structure throughout, neglecting possible exogenous or endogenous cost efficiency effects. Noting that because of stationary strategies we can write $TAC_i(q_i) = AC_{it}(q_i) \cdot \sum_{t=2011}^{2110}(1+r)^{-(t-2010)}$ and discounting abatement costs with the same uniform discount rate of 2 percent as in the case of benefits, we get $TAC_i(q_i) = 43.1 \cdot AC_{it}(q_i)$ and marginal total abatement costs of $MTAC_i(q_i) = 43.1 \cdot MAC_{it}(q_i)$ which are drawn in Figure 1.

Figure 1: Marginal Total Abatement Cost Functions



From the graph it is evident that marginal abatement cost never intersect and that CHN and USA have the flattest curves and BRA and JPN the steepest.

3.6 Payoff Function

Using the information of section 3.4 and 3.5 gives the following payoff function:

[14] $\pi_i = TB_i(q) - TAC_i(q_i)$

Since we consider 12 world regions, this gives rise to 4096 different membership strategy vectors in the first stage of the coalition formation game. However, since a strategy vector where only one region announces $\sigma_i = 1$ and all other regions announce $\sigma_i = 0$ leads to the same coalition structure as if all regions announce $\sigma_i = 0$, the set of coalition structures, C, comprises "only" 4084 different coalition structures. For each coalition structure we compute a payoff vector according to the assumption of the second stage of the coalition formation game (see Definition 2) that coalition members jointly maximize the aggregate payoff to their coalition. For coalition structure c, $\sum_{i \in c^i} MTB_i(q) = MTAC_i(q_i)$ holds for coalition c^i in equilibrium. Since our specification of $TB_i(q)$ implies a linear function and hence constant marginal benefits, signatories and non-signatories have dominant abatement strategies. That is, optimal abatement strategies of a region or group of regions are independent of those of other regions. This implies that if regions form a coalition, and thereby increasing their abatement efforts, this is not offset by a reduction of abatement efforts by outsiders. In other words, in our model no leakage effects occur. According to theory (Carraro/Siniscalco 1998 and Finus 2003a), this is the most favorable condition for forming stable coalitions. Nevertheless, as will be apparent from subsequent sections, cooperation proves very difficult.

Finally note that in the context of our empirical model, individual rationality is a necessary condition for internal stability of a coalition and hence a necessary condition under open and exclusive membership. Individual rationality implies that all signatories must receive a higher payoff than in the singleton coalition structure. Hence, if this condition is violated for at least one signatory, we can immediately conclude that this coalition cannot be stable, regardless whether we apply Definition 3, 4 or 5.[8]

4. Results of the Base Case

4.1 Introduction

From the previous discussion it is evident that in particular the estimation of benefits from global abatement is associated with some uncertainty. This concerns primarily the level of damages represented by the parameter γ_D. Hence, we discuss first the "base case" value of

---

[8] A proof is available upon request from the authors.

$\gamma_D = 0.027$ in section 4 and then conduct a sensitivity analysis in section 5 where we consider different levels of this value.

In order to gain insight in the fundamental features of our model, we discuss first three benchmark scenarios in subsection 4.2: 1) the *singleton coalition structure*, implying no cooperation. 2) The *grand coalition*, implying full cooperation. 3) The *Kyoto coalition* that constitutes partial cooperation. Here we assume that the members of the original Kyoto Protocol (before the USA withdrew from the Protocol) form a coalition, which includes USA, JPN, EEC, OOE, EET and FSU. Subsequently, we report on results of our stability check (subsection 4.3).

4.2 Benchmark Scenarios

a) Singleton Coalition Structure

Table 2 reports results if each region acts by itself, corresponding to the "classical" Nash equilibrium with no cooperation. In equilibrium, marginal abatement costs are equal to marginal benefits for each region. Annual global emission reduction amounts to only 4.6 percent. This implies a global stock of CO_2 of 1,561 gigatons in 2110. This is about 2.5 times of the pre-industrial level. The fact that benefits are rather high compared to costs from abatement explains that even in the absence of any cooperation total emission reduction exceeds that in the BAU-scenario (no abatement) by 55 gigatons.

Table 2: Singleton Coalition Structure (Nash Equilibrium)

Regions	Total emission reduction	Annual emission reduction	Total abatement costs	Total benefits from abatement	Benefits minus abatement costs	Marginal abatement costs	Marginal benefits
	Gton (over 100 years)	percentage of emissions in 2010	bln US$ over 100 years	bln US$ over 100 years	bln US$ over 100 years	US$/ton	US$/ton
USA	16	6.7	53	468	415	8.5	8.5
JPN	1	1.4	2	357	354	6.5	6.5
EEC	7	4.7	24	488	464	8.8	8.8
OOE	2	3.1	1	71	71	1.3	1.3
EET	1	1.8	0	27	27	0.5	0.5
FSU	5	4.9	4	140	135	2.5	2.5
EEX	1	0.7	0	62	62	1.1	1.1
CHN	15	6.6	16	128	112	2.3	2.3
IND	3	5.3	3	103	101	1.9	1.9
DAE	1	1.3	0	52	51	0.9	0.9
BRA	0	0.1	0	32	32	0.6	0.6
ROW	4	5.3	4	141	137	2.5	2.5
World	55	4.6	109	2,069	1,960		

Global stock of carbon dioxide by 2110 = 1,561 Gton

At the level of individual regions, it is evident that annual emission reductions vary widely. The reason is large differences in marginal abatement cost curves (see Figure 1 and Table 1, section 3) and marginal benefits from abatement (see Table 1, section 3) between regions. For instance, USA has a relatively flat marginal abatement cost curve. Hence, in the absence of cooperation, USA has an incentive to reduce annual emissions by 6.7 percentage because of her high marginal benefits from abatement. A similar argument applies to CHN that has an even flatter marginal abatement cost curve, though lower marginal benefits from abatement compared to USA. In contrast, regions like BRA and DAE have virtually no incentive at all to conduct emission reductions by themselves because of steep marginal abatement cost curves and low marginal benefits from abatement.

b) Grand Coalition

Table 3 displays results for the grand coalition that corresponds to the "classical" global optimum with full cooperation. In equilibrium, marginal abatement costs are equal across countries and amount to 37.4 US$/ton - a value that is in the range of many other empirical studies (e.g., Weyant 1999). At the aggregate level, the reduction of annual emission reductions amount to 21.4 percent, exceeding those in the singleton coalition structure by a substantial amount. Nevertheless, the effect on the final stock of CO_2 is only moderate - a feature reminiscent also to most computable general equilibrium models: it amounts to a reduction of only 5.5 percentage compared to the singleton case and 6.9 percent compared to BAU. The reason is that the airborne fraction of CO_2-emissions that remains in the atmosphere is only 64 percent and the annual natural removal rate of 0.86 percent levels off differences between both scenarios over a time period of 100 years. However, the gains from cooperation are considerable: the total payoff (benefits minus abatement costs) in the grand coalition is 6,031 billion US$, which implies a gain of 208 percent compared to the singleton coalition structure. This figure highlights the relevance of cooperation in the case of climate change.

Table 3: Grand Coalition (Global Optimum)*

Regions	Total emission reduction	Annual emission reduction	Total abatement costs	Total benefits from abatement	Benefits minus abatement costs	Marginal abatement costs	Marginal benefits	Incentive to leave coalition
	Gton (over 100 years)	percentage of emissions in 2010	bln US$ over 100 years	Bln US$ over 100 years	bln US$ over 100 years	US$/ton	US$/ton	bln US$ over 100 years
USA	38	15.7	513	2,169	1,656	37.4	8.5	23.6
JPN	4	6.5	63	1,653	1,590	37.4	6.5	-123.8
EEC	16	11.5	229	2,262	2,033	37.4	8.8	-180.1
OOE	10	16.5	127	331	203	37.4	1.3	109.6
EET	10	19.6	130	125	**-6**	37.4	0.5	124.9
FSU	19	19.3	242	647	405	37.4	2.5	178.1
EEX	12	10.2	188	288	99	37.4	1.1	169.9
CHN	96	40.6	1,348	594	**-754**	37.4	2.3	1133.2
IND	22	33.8	295	479	184	37.4	1.9	245.8
DAE	10	25.1	155	239	84	37.4	0.9	142.1
BRA	1	5.5	12	147	135	37.4	0.6	10.0
ROW	19	26.5	250	652	401	37.4	2.5	185.1
World	256	21.4	3,553	9,584	6,031			-

Global stock of carbon dioxide by 2110 = 1,475 Gton

* Last column: gain from leaving coalition if all other regions remain in coalition.

A closer inspection of individual regions reveals that, in terms of total emission reduction, CHN, USA and IND have to contribute more than other regions to a globally optimal solution due to their flat marginal abatement cost curves. For EET and CHN a globally optimal solution would not be individually rational since these regions would loose compared to the Nash equilibrium as is indicated by bold faced figures in Table 3, column 6. Those regions have to contribute much to cooperation but benefit only little in the form of reduced damages. Therefore, it is clear that the grand coalition is not stable. Moreover, a more detailed analysis considering the last column of Table 3 reveals that all regions, except JPN and EEC, have an incentive to leave the grand coalition. Considering the absolute amount of the gains from leaving the grand coalition suggests that most regions face a strong free-rider incentive.

c) Kyoto Coalition

Table 4 displays results for the Kyoto coalition. Hence, according to the assumption of the valuation function, the first 6 regions (indicated in italics in Table 4) jointly maximize the aggregate payoff to their coalition and therefore marginal abatement costs of these regions are equal. The annual global emission reduction is substantially lower than in the global optimum but almost twice as high as in the Nash equilibrium. Also the global gain from cooperation is with 3,140 bln US$, 60 percent higher than in the Nash equilibrium.

Table 4: Kyoto Coalition*

Regions	Total emission reduction	Annual emission reduction	Total abatement costs	Total benefits from abatement	Benefits minus abatement costs	Marginal abatement costs	Marginal benefits	Incentive to change membership strategy
	Gton (over 100 years)	percentage of emissions in 2010	bln US$ over 100 years	bln US$ over 100 years	bln US$ over 100 years	US$/ton	US$/ton	bln US$ over 100 years
USA	32	13.4	332	906	574	28.0	8.5	65.3
JPN	3	5.2	38	691	653	28.0	6.5	-46.9
EEC	14	9.7	147	945	798	28.0	8.8	-52.8
OOE	9	14.3	83	138	**55**	28.0	1.3	70.5
EET	9	16.9	85	52	**-33**	28.0	0.5	80.3
FSU	17	16.7	157	270	**113**	28.0	2.5	114.6
EEX	1	0.7	0	120	120	1.1	1.1	-113.5
CHN	15	6.6	16	248	232	2.3	2.3	-794.9
IND	3	5.3	3	200	197	1.9	1.9	-172.7
DAE	1	1.3	0	100	99	0.9	0.9	-93.9
BRA	0	0.1	0	61	61	0.6	0.6	-6.5
ROW	4	5.3	4	272	268	2.5	2.5	-137.8
World	107	8.9	865	4,005	3,140			

Global stock of carbon dioxide by 2110 = 1,539 Gton

* Last column: gain from changing membership strategy if all other regions stick to their membership strategy.

However, regardless of the membership rule, the Kyoto coalition is not stable since internal stability is violated. First, individual rationality is violated for the three coalition members OOE, EET and FSU that would be worse off than in the singleton coalition structure as indicated by bold faced numbers in Table 4, column 5. Second, not only these regions but also the USA have an incentive to leave the coalition as is evident from the last column in Table 4. This result together with our finding that the USA will already conduct relative high abatement without cooperation (see Table 2) helps to explain the decision of President Bush to withdraw from the Kyoto Protocol and his announcement to pursue, nevertheless, an "active" national climate policy.[9]

Not surprising, all 6 outsiders are better off than in the Nash equilibrium since they benefit from the abatement efforts of the Kyoto coalition. More surprising is the fact that none of the outsiders has an incentive to join the coalition, which follows from the negative number in the last column in Table 4. The reason is that if already 6 regions have formed a coalition, joining would imply a substantial increase of abatement efforts for a potential entrant but only a marginal additional benefit from reduced emissions.

---

[9] It should be pointed out that our analysis does not consider permit trading. However, at the time of the withdrawal of the USA, it was still under discussion whether a permit trading system will be established at all and if whether trade will be unrestricted.

4.3 Stability Analysis

We checked all 4084 coalition structures for stability under open membership, exclusive membership with majority voting and exclusive membership with unanimity voting with an algorithm programmed with the software package Matlab. We found no non-trivial coalition structure that is stable under open membership.[10] Whereas more than 1000 coalition structures are externally stable (under open membership), only 14 coalition structures are internally stable. Thus, it seems that main problem of cooperation is internal instability due the presence of strong free-rider incentives to leave a coalition. Nevertheless, under exclusive membership and majority voting at least one out of the 14 internally stable coalitions is externally stable of which the results are displayed in Table 5.

Table 5: Coalition between FSU, BRA and ROW*

Regions	Total emission reduction	Annual emission reduction	Total abatement costs	Total benefits from abatement	Benefits minus abatement costs	Marginal abatement costs	Marginal benefits	Incentive to change membership strategy
	Gton (over 100 years)	percentage of emissions in 2010	bln US$ over 100 years	bln US$ over 100 years	bln US$ over 100 years	US$/ton	US$/ton	bln US$ over 100 years
USA	16	6.7	53	511	458	8.5	8.5	61.4
JPN	1	1.4	2	390	387	6.5	6.5	45.0
EEC	7	4.7	24	533	509	8.8	8.8	76.1
OOE	2	3.1	1	78	77	1.3	1.3	-4.0
EET	1	1.8	0	29	29	0.5	0.5	-6.6
FSU	7	7.4	14	153	138	5.6	2.5	-1.4
EEX	1	0.7	0	68	67	1.1	1.1	-6.0
CHN	15	6.6	16	140	124	2.3	2.3	-49.0
IND	3	5.3	3	113	110	1.9	1.9	-8.5
DAE	1	1.3	0	56	56	0.9	0.9	-5.2
BRA	0	1.1	0	35	34	5.6	0.6	-0.1
ROW	6	8.9	14	154	140	5.6	2.5	-1.5
World	60	5.1	129	2,260	2,131	-	-	-

Global stock of carbon dioxide by 2110 = 1,559 Gton

* Stable under exclusive membership and majority and unanimity voting. Last column: gain from changing membership strategy if all other regions stick to their membership strategy.

From Table 5 it is evident that internal stability holds "at the margin". Each of the three members (indicated in italics) would only slightly loose by leaving the coalition as shown in the last column. Moreover, each member only slightly gains compared to the singleton coalition structure (see Table 2). Of course, non-signatories gain more since they benefit from additional abatement efforts of the coalition without carrying additional abatement costs.

---

[10] A non-trivial coalition structure includes a coalition with at least two members. In the following, we concentrate in the stability analysis on these coalition structures since the singleton coalition structure is stable by definition. See section 2.

Overall, this coalition only marginally improves upon the singleton case. Global net benefits are 2,131 billion US$ whereas in the singleton coalition structure they are 1,960 billion US$. This implies only an improvement of 8.7 percent, a rather modest contribution to tackle global warming given that the global optimum would imply an improvement of almost 208 percent. In other words, stable partial cooperation can only slightly reduce the gap between the Nash equilibrium and the global optimum - a phenomenon confirmed by all other results discussed below. The reason is closely related to the fundamental features of coalition formation.

Internal stability will only hold for coalitions that slightly increase their abatement efforts compared to the singleton coalition structure. If abatement efforts were more ambitious, the free-rider incentive would become too strong so that internal stability would fail to hold. For instance, in the context of a coalition between FSU, BRA and ROW, all three members have relatively high marginal abatement costs compared to marginal benefits from abatement. Hence, these regions increase their abatement efforts only slightly compared to the singleton coalition structure. Consequently, annual global emission reduction is 5.1 percent, which is only slightly higher than in the singleton coalition structure, which amounts to 4.6 percent. Thus, the conjecture would be wrong that stable coalitions are formed by the "good guys" and that outsiders are the "bad guys". For instance, BRA´s annual emission reduction is 1.1 percent in this coalition whereas the outsiders USA, EEC and CHN reduce their emissions on average by 6.7, 4.7 and 6.6 percent, respectively.

In terms of *external stability* it is evident that only USA, JPN and EEC have an incentive to join the coalition formed by FSU, BRA and ROW. The reason is that the potential entrants would receive a large portion of the gains from joint cooperation: these regions have high marginal benefits from global abatement compared to their contribution to joint abatement in a coalition (because of relatively steep marginal abatement cost curves). It is exactly for this reason why their application for accession is turned down by a majority of regions in the coalition.

The explanations about internal stability also help to rationalize participation in stable coalitions. Only coalitions including members that agree on low abatement targets and where members exhibit a similar cost-benefit structure are stable (e.g., FSU, BRA, ROW). High abatement targets would imply a high free-rider incentive. A heterogeneous cost-benefit structure (e.g., including additionally USA, JPN or EEC) would imply an asymmetric distribution of the gains from cooperation, putting some countries at disadvantage, which have therefore an incentive to leave the coalition.

Of course, by definition, the coalition between FSU, BRA and ROW is also stable under unanimity voting. However, the following two coalition structures, listed in Table 6 and 7, are only stable under exclusive membership and unanimity voting, but not under majority voting.

Table 6: Coalition between CHN and DAE*

Regions	Total emission reduction	Annual emission reduction	Total abatement costs	Total benefits from abatement	Benefits minus abatement costs	Marginal abatement costs	Marginal benefits	Incentive to change membership strategy
	Gton (over 100 years)	percentage of emissions in 2010	Bln US$ over 100 years	bln US$ over 100 years	bln US$ over 100 years	US$/ton	US$/ton	bln US$ over 100 years
USA	16	6.7	53	515	462	8.5	8.5	253.0
JPN	1	1.4	2	393	390	6.5	6.5	157.8
EEC	7	4.7	24	537	513	8.8	8.8	276.2
OOE	2	3.1	1	79	78	1.3	1.3	5.2
EET	1	1.8	0	30	29	0.5	0.5	-1.8
FSU	5	4.9	4	154	149	2.5	2.5	23.8
EEX	1	0.7	0	68	68	1.1	1.1	3.0
CHN	20	8.5	28	141	113	3.2	2.3	-0.6
IND	3	5.3	3	114	111	1.9	1.9	10.7
DAE	2	4.1	3	57	54	3.2	0.9	-2.9
BRA	0	0.1	0	35	35	0.6	0.6	1.4
ROW	4	5.3	4	155	151	2.5	2.5	24.1
World	61	5.1	123	2,277	2,153	-	-	-

Global stock of carbon dioxide by 2110 = 1,559 Gton

* Stable under exclusive membership and unanimity voting. Last column: gain from changing membership strategy if all other regions stick to their membership strategy.

Table 7: Coalition between EEX and CHN*

Regions	Total emission reduction	Annual emission reduction	Total abatement costs	Total benefits from abatement	Benefits minus abatement costs	Marginal abatement costs	Marginal benefits	Incentive to change membership strategy
	Gton (over 100 years)	percentage of emissions in 2010	bln US$ over 100 years	bln US$ over 100 years	bln US$ over 100 years	US$/ton	US$/ton	bln US$ over 100 years
USA	16	6.7	53	524	471	8.5	8.5	254.6
JPN	1	1.4	2	400	397	6.5	6.5	158.9
EEC	7	4.7	24	547	523	8.8	8.8	278.4
OOE	2	3.1	1	80	79	1.3	1.3	5.0
EET	1	1.8	0	30	30	0.5	0.5	-2.0
FSU	5	4.9	4	156	152	2.5	2.5	23.5
EEX	2	1.8	3	70	66	3.4	1.1	-4.4
CHN	21	8.8	31	144	112	3.4	2.3	-0.3
IND	3	5.3	3	116	113	1.9	1.9	10.2
DAE	1	1.3	0	58	57	0.9	0.9	1.5
BRA	0	0.1	0	35	35	0.6	0.6	1.4
ROW	4	5.3	4	158	154	2.5	2.5	23.8
World	62	5.2	127	2,317	2,190	-	-	-

Global stock of carbon dioxide by 2110 = 1,558 Gton

* Stable under exclusive membership and unanimity voting. Last column: gain from changing membership strategy if all other regions stick to their membership strategy.

From Table 6 and 7 it is evident that most observations and conclusions from above also apply for a coalition between CHN and DAE and a coalition between EEX and CHN.[11] First, internal stability for coalition members only holds at the margin. Second, both coalitions only marginal improve upon the singleton coalition structure in terms of abatement and global net benefits because stable coalitions comprise members that only slightly increase their abatement efforts compared to the Nash equilibrium. Third, USA, JPN and EEC are those regions with the strongest incentive to join the coalition which can only be "neutralized" by unanimity vote.

Viewed together, three important conclusions emerge from the stability analysis of the base case. *First*, not surprisingly, our theoretical prediction $C(OM) \subset C(EM\text{-}M) \subset C(EM\text{-}U)$ is confirmed in our empirical analysis. Whereas no non-trivial coalition is stable under open membership, one non-trivial coalition is stable under exclusive membership and majority voting and two other additional coalitions are stable under unanimity voting. However, a more interesting fact is that those additional coalitions are superior in net benefit and ecological terms. The coalition between FSU, BRA and ROW, which is only stable under exclusive membership but not under open membership, implies total abatement of 60 gigatons CO_2 over 100 years and total net benefits of 2,131 billion US$. In contrast, in the Nash equilibrium this is 55 gigatons CO_2 and 1,960 billion US$. Moreover, additional coalitions that are only stable under exclusive membership and unanimity voting (but not majority voting) constitute a further improvement. The coalition between CHN and DAE implies a total emission reduction of 61 gigatons CO_2 and global net benefits of 2,153 billion US$. The coalition between EEX and CHN implies a total emission reduction of 62 gigatons CO_2 and global net benefits of 2,190 billion US$. Hence, exclusive membership leads to superior outcomes than open membership and unanimity voting leads to superior outcomes than majority voting.

Second, stable coalition that constitute partial cooperation only marginally close the gap between the Nash equilibrium and the global optimum if this gap is large. As argued above, in our context the gap in net benefits is large in absolute and relative terms since global net benefits in the global optimum are 6,031 billion US$ and in the Nash equilibrium 1,960 billion US$. This implies that net benefits are 208 percent higher in the global optimum than in the Nash equilibrium. In contrast, the coalition between FSU, BRA and ROW, CHN and DAE

---

[11] It does not seem evident which of the two coalitions will form. On the one hand, CHN has a slight preference to form a coalition with DAE instead of with EEX. On the other hand, DAE prefers a coalition between CHN and EEX and EEX prefers a coalition between CHN and DAE.

and EEX and CHN imply only an improvement in net benefits of 8.7, 9.8 and 11.7 percent, respectively.

Third, despite the fact that the coalition between FSU, BRA and ROW counts one more member than the coalition between CHN and DAE and the coalition between EEX and CHN, it is inferior in terms of global net benefits and abatement. This indicates that success of an IEA cannot be inferred from the number of participants. All three conclusions will be confirmed by our sensitivity analysis on which we report in section 5 and where we argue that they are perfectly in line with results obtained by theory.

5 Sensitivity Analysis

A typical feature of empirical work is that results depend on parameter values, which are subject to some uncertainty. Given the large number of parameters that enter our model, some selection is necessary for a sensitivity analysis. We believe that the highest uncertainty concerns benefits from global abatement. Hence, we conduct a sensitivity analysis where we uniformly lower or raise the level of benefits from global abatement. That is, we change the base value of $\gamma_D = 0.027$. For instance, lowering global benefits to 50 percent compared to the base case implies $\gamma_D = 0.014$, which is very close to the value estimated by Nordhaus (1994). Raising this value to 120, 200 and 300 percent implies higher benefits compared to our base case value of $\gamma_D = 0.027$ (100 percent) taken from Tol (1997). Table 8 summarizes the results for five scenarios: 50, 100, 120, 200 and 300 percent of $\gamma_D = 0.027$.

First, not only for our base case but also for all scenarios neither the grand coalition structure nor the Kyoto coalition structure is stable regardless of the definition of stability. Although both coalition structures would substantially improve upon the singleton coalition structure, some members have a strong incentive to leave the coalition.

Second, partial cooperation is not stable under open membership for the 50 percent and 100 percent scenario. Only if we raise benefits sufficiently high, one coalition between JPN and EEC is stable under open membership. It turns out that the lower benchmark is 120 percent. However, above this level, no additional coalition is stable under open membership.

Also under exclusive membership stability rises with the level of benefits from global abatement. For instance, in the 50 percent scenario only one non-trivial coalition structure is stable under exclusive membership and unanimity voting, three in the 100 percent scenario and four in the 120, 200 and 300 percent scenarios. The reason is that higher benefits imply that regions already conduct higher abatement in the Nash equilibrium so that partial cooperation requires only small additional abatement efforts. For instance, considering the coalition between FSU, BRA, and ROW we find in the 50 percent scenario that total emission reduction exceeds that in the singleton coalition structure by 11.7 percent, in the 100 percent scenario by 9.1 percent and in the 300 percent scenario by only 8.0 percent. Hence, the free-rider incentive in terms of internal stability decreases with the level of benefits.

Table 8: Sensitivity Analysis*

Scenario	Coalitions	Stability			Total emission reduction	Annual emission reduction	Total abatement costs	Total benefits from abatement	Benefits minus abatement costs
		OM (1)	EM (2)		Gton (over 100 years)	percentage of emissions in 2010	bln US$ over 100 years	bln US$ over 100 years	bln US$ over 100 years
			M (3)	U (4)					
50 %	singleton coalition structure	X	X	X	34	2.9	36	644	608
	Coalition JPN, EEC	–	–	–	39	3.3	64	736	672
	Coalition FSU, BRA, ROW	–	X	X	38	3.2	43	709	667
	Coalition CHN, DAE	–	–	–	38	3.2	40	708	668
	Coalition EEX, CHN	–	–	–	39	3.2	41	721	679
	Kyoto coalition	–	–	–	70	5.9	298	1,317	1,018
	grand coalition	–	–	–	172	14.4	1,225	3,211	1,986
100 % base case	singleton coalition structure	X	X	X	55	4.6	109	2,069	1,960
	Coalition JPN, EEC	–	–	–	59	5.0	152	2,208	2,056
	Coalition FSU, BRA, ROW	–	X	X	60	5.1	129	2,260	2,131
	Coalition CHN, DAE	–	–	X	61	5.1	123	2,277	2,153
	Coalition EEX, CHN	–	–	X	62	5.2	127	2,317	2,190
	Kyoto coalition	–	–	–	107	8.9	865	4,005	3,140
	grand coalition	–	–	–	256	21.4	3,553	9,584	6,031
120 %	singleton coalition structure	X	X	X	62	5.2	145	2,801	2,655
	Coalition JPN, EEC	X	X	X	67	5.6	203	2,988	2,784
	Coalition FSU, BRA, ROW	–	X	X	68	5.7	172	3,053	2,881
	Coalition CHN, DAE	–	–	X	69	5.7	165	3,081	2,917
	Coalition EEX, CHN	–	–	X	70	5.8	170	3,136	2,966
	Kyoto coalition	–	–	–	119	10.0	1,143	5,354	4,211
	grand coalition	–	–	–	284	23.8	4693	12,746	8,053
200 %	singleton coalition structure	X	X	X	87	7.3	324	6,485	6,161
	Coalition JPN, EEC	X	X	X	92	7.7	455	6,908	6,453
	Coalition FSU, BRA, ROW	–	X	X	94	7.9	382	7,035	6,652
	Coalition CHN, DAE	–	–	X	95	8.0	369	7,128	6,760
	Coalition EEX, CHN	–	–	X	97	8.1	380	7,251	6,871
	Kyoto coalition	–	–	–	161	13.4	2,490	12,025	9,535
	grand coalition	–	–	–	377	31.5	10,204	28,205	18,000
300 %	singleton coalition structure	X	X	X	112	9.3	609	12,519	11,910
	Coalition JPN, EEC	X	X	X	119	9.9	857	13,323	12,466
	Coalition FSU, BRA, ROW	–	X	X	121	10.1	716	13,537	12,821
	Coalition CHN, DAE	–	–	X	123	10.2	694	13,746	13,052
	Coalition EEX, CHN	–	–	X	125	10.4	714	13,976	13,262
	Kyoto coalition	–	–	–	203	17.0	4,610	22,751	18,142
	grand coalition	–	–	–	470	39.3	18,856	52,759	33,903

* (1) OM = open membership, (2) EM= exclusive membership, (3) M = majority voting, (4) U = unanimity voting; X means stable and – means not stable; scenarios imply “percentage of the benefit parameter $\gamma_D = 0.027$ ”.

Third, and closely related to the second point, a conclusion that has been derived from theoretical models under very restrictive assumptions (symmetric countries or heterogeneous countries with only two types of countries) and that has been called a paradox by Barrett (1994 and 1997) is confirmed: *whenever cooperation would be needed most from a global point of view* (the relative gap between Nash equilibrium and the global optimum) *stable partial cooperation achieves only little*. To see this, we compute in a first step the relative difference in net benefits between the singleton coalition structure and the grand coalition. It turns out that this difference gradually decreases from 226.6 percent in the 50 percent scenario to 184 percent in the 300 percent scenario. That is, the degree of externality decreases with the level of benefits (and fixed abatement cost parameters) because regions will already unilaterally conduct substantial abatement efforts in the Nash equilibrium if benefits are high. In a second step, we compute the relative difference in net benefits between stable coalitions and the singleton coalition structure. For instance, under exclusive membership with unanimity voting only a coalition between FSU, BRA and ROW is stable in the 50 percent scenario that improves upon the Nash equilibrium by 9.7 percent. In contrast, in the 300 percent scenario the highest global net benefits are generated by the coalition between EEX and CHN with an improvement of 11.35 percent. Similar, under open membership, no coalition is stable in the 50 and 100 percent scenario and in the 120, 200 and 300 percent scenario a coalition of JPN and EEC is stable which improves upon the Nash equilibrium by roughly 4.7 percent. Thus, it is evident that when the gap between the Nash equilibrium and the global optimum is large, either partial cooperation is not stable, or if it is stable, it closes the gap only by a small amount. In contrast, if the gap is smaller, partial cooperation is more successful.

Fourth, independent of the stability concept, if there are stable coalitions, they are rather small. However, the number of participants is no indication of the success of cooperation. Not only for the base case is the coalition of three members formed by FSU, BRA and ROW inferior to the coalition between CHN and DAE and EEX and CHN, each comprising only two members, but also in the scenarios 120, 200 and 300 percent. Again, this result is line with findings of theory (see Finus 2001 and 2003a for details).

Fifth, not only for the base case but also for the other scenarios it is confirmed that exclusive membership leads to superior outcomes than open membership and that under exclusive membership unanimity voting leads to superior outcomes than majority voting.

6. Summary and Conclusions

We studied stability of climate change coalitions. We followed the standard assumption of the valuation function approach that coalition members efficiently choose their abatement levels as to jointly maximize aggregate net benefits to their coalition. We departed from the standard assumption of the definition of internal&external stability that implies open membership and also considered exclusive membership with two voting schemes: majority voting and unanimity voting. We applied these different notions of stability to an empirical model, called STACO. This model captures important dynamic properties of the global warming problem but assumes stationary abatement strategies for game theoretic tractability. It covers the period between 2010 and 2110 and comprises 12 world regions, giving rise to 4084 different coalition structures. Thus, though this model is far from being perfect, it improves upon previous studies in that it captures more dynamic aspects of the climate change problem and allows for a more detailed analysis of the strategic interaction of many actors. We conducted several sensitivity analyses (called scenarios) in order to test the robustness of our results.

First, in the case of global warming, the gains from cooperation prove to be large in our model. This does not only hold in absolute but also in relative terms when global net benefits in the global optimum are compared with those in the Nash equilibrium. This stresses the importance of economic and game theoretic analyses on global warming for future research.

Second, neither the grand coalition nor the Kyoto coalition (comprising the countries of the original Kyoto Protocol before the USA withdrew from this agreement) turned out to be stable for all scenarios regardless of the membership rule and voting scheme, though both coalition structures would substantial improve upon the Nash equilibrium. In contrast, those coalitions that turned out to be stable only comprise few members and only marginally improve upon the Nash equilibrium in terms of global net benefits, global emissions and stock of greenhouse gases. This result provided some rationale why it proves so difficult to achieve a high participation in the Kyoto Protocol and to agree on ambitious and effective abatement targets. The more countries accede to an agreement and/or the higher are abatement targets, the more difficult it becomes to control free-riding. Thus, as long as no major changes of the underlying economic incentive structure occur, it seems that only moderate progress can be expected in the near future in the context of climate change. However, it has to be pointed out that our pessimistic conclusion is based on two assumptions. Dropping these assumptions is closely related to possible policy measures that may foster stable and effective cooperation and that we would like to analyze in future research.

One option is to drop the assumption of no transfers and include various transfer schemes in the stability analysis. This may comprise direct monetary transfers as suggested by the meeting of parties to the Kyoto Protocol in Marrakech. The proposal allows developing countries to draw on financial resources from an environmental fund, as in the case of the Montreal Protocol. However, transfers may also comprise indirect measures as for instance permit trading (Article 17), clean development mechanism (Article 12) and joint implementation (Articles 3 and 4). We suspect that all kind of transfers will help to balance different interests between countries (see, e.g., Botteon/Carraro 1997 and Barrett 1997). In particular, we expect a higher participation, which also allows to reap higher efficiency gains from cooperation because stable coalitions can be formed by more heterogeneous countries.

Another option is to drop the assumption of joint welfare maximization that implies not only that cost efficient but also ambitious abatement targets are implemented within coalitions. This is one important reason for instability of large coalitions because of high free-rider incentives and an unequal distribution of the gains from cooperation. Thus, it is likely that better results may be achieved if members settle for less ambitious abatement targets and/or if abatement burdens are allocated more equally (though cost-inefficiently). If the effect on participation is strong enough, this may well compensate for inefficiencies. For instance, theoretical work by Endres/Finus (2002), Finus/Rundshagen (1998) and Finus (2003b) suggest that there may some rationale that within many IEAs abatement targets are specified as inefficient uniform emission reduction quotas and that signatories agree on the lowest common denominator (resulting from unanimity voting) in terms of the joint abatement level. These extensions could be further steps in including public choice aspects in the analysis of international treaty formation that may help to rationalize inefficient designs of many actual IEAs.

Third, we confirmed a theoretical result that has been derived under very restrictive assumption: whenever the degree of externality - measured as the relative difference between the Nash equilibrium (singleton coalition structure) and the global optimum (grand coalition) - is large, partial cooperation is either not stable or achieves only little. Only if this difference is small enough, more progress can be expected. In our model, this difference became smaller when we (uniformly) raised benefits from global abatement. Then regions already conduct a substantial amount of abatement without any cooperation, additional abatement requirements within a coalition are relatively small and hence free-riding is less of a problem. From a policy point of view, this suggests that the success of mitigating the global warming problem is closely related to the perception of environmental damages of governments and societies. Thus, measures that enhance environmental consciousness may foster cooperation in the future.

Fourth, large stable coalitions may prove to be inferior compared to small stable coalitions both in economic and ecological terms. This suggests that a high participation in an IEA does not necessarily imply its success, as frequently publicized by politicians and the media. Only a comparison of net benefits of an agreement with an appropriate benchmark (Nash equilibrium or business as usual scenario) allows to draw sound conclusions.

Fifth, stability proves generally very difficult under open membership rule. We find only for some of our scenarios one coalition that is stable under open membership. In contrast, some of the coalitions that are externally unstable under open membership turn out to be stable under exclusive membership. Those additional stable coalitions under exclusive membership generate higher global net benefits and a lower stock of greenhouse gases. It turns out that unanimity voting leads to better outcomes than majority voting.

From a theoretical point of view this suggests not only that a modification of the stability concept of internal&external stability is a fruitful route for research, but also that previous results may have been overly pessimistic. From an applied point of view, the results are interesting in two respects. First, almost all past IEAs have no provision to restrict membership. Hence, it may be worthwhile to think whether to adopt an exclusive membership rule, which is typical for club good agreements, as for instance NATO and European Union, also for public good agreements like those on climate change in the future. Second, exclusive membership requires some degree of consensus between coalition partners to form a coalition. At first glance, intuition suggests that the higher the degree of consensus needed to come to a decision, the less likely and successful will be an agreement. However, a closer inspection reveals that when the free-rider problem is explicitly accounted for this conclusion has to be questioned. This qualification is confirmed in the literature that analyzes bargaining on environmental policy levels (e.g., Endres 1997, Endres/Finus 2002 and Finus/Rundshagen 1998) and is also confirmed by our model that analyzes bargaining on membership. In both cases, consensus agreements are associated with more stability that is a basic prerequisite for any successful agreement. Hence, our results provide a further rationale for the frequent application of consensus voting and in particular for unanimity voting as applied in many international organizations and agreements.

References

Babiker, M.H., J.M. Reilly, M. Mayer, R.S. Eckaus, I.S. Wing, and R.C. Hyman (2001), The MIT Emissions Prediction and Policy Analysis (EPPA) Model: Revisions, Sensitivities, and Comparison Results, Report No. 71, MIT, Cambridge.

Barrett, S. (1994), Self-Enforcing International Environmental Agreements. "Oxford Economic Papers", vol. 46, pp. 804-878.

Barrett, S. (1997), Heterogeneous International Agreements. In: C. Carraro (ed.), International Environmental Negotiations: Strategic Policy Issues. Edward Elgar, Cheltenham, pp. 9-25.

Bauer, A. (1992), International Cooperation over Greenhouse Gas Abatement. Mimeo, Seminar für empirische Wirtschaftsforschung, University of Munich, Munich.

Bloch, F. (1997), Non-Cooperative Models of Coalition Formation in Games with Spillovers. In: Carraro, C. and D. Siniscalco (eds.), New Directions in the Economic Theory of the Environment. Cambridge University Press, Cambridge, ch. 10, pp. 311-352.

Böhringer, C. and C. Vogt (2003), The Dismantling of a Breakthrough: The Kyoto Protocol - Just Symbolic Policy! ZEW-Discussion Paper, No. 02-25, Mannheim.

Botteon, M. and C. Carraro (1997), Burden-Sharing and Coalition Stability in Environmental Negotiations with Asymmetric Countries. In: Carraro, C. (ed.), International Environmental Negotiations: Strategic Policy Issues. Edward Elgar, Cheltenham, ch. 3, pp. 26-55.

Buchanan, J. M. and G. Tullock (1962), The Calculus of Consent. University of Michigan Press, Ann Arbour.

Carraro, C. and D. Siniscalco (1993), Strategies for the International Protection of the Environment. "Journal of Public Economics", vol. 52, pp. 309-328.

Carraro, C. and D. Siniscalco (1998), International Environmental Agreements: Incentives and Political Economy. "European Economic Review", vol. 42, pp. 561-572.

Chander, P. and H. Tulkens (1995), A Core-Theoretic Solution for the Design of Cooperative Agreements on Transfrontier Pollution. "International Tax and Public Finance", vol. 2, pp. 279-293.

Chander, P. and H. Tulkens (1997), The Core of an Economy with Multilateral Environmental Externalities. "International Journal of Game Theory", vol. 26, pp. 379-401.

Dellink, R.B., M. Finus, E.C. van Ierland and J.-C. Altamirano (2003), Empirical Background Paper of the STACO Model, Draft, University of Wageningen.

Dijkstra, B. R. (1999), The Political Economy of Environmental Policy: A Public Choice Approach to Market Instruments. Edward Elgar, Cheltenham, UK.

Ellerman, A.D. and A. Decaux (1998), MIT Report #40. Analysis of Post-Kyoto CO_2 Emissions Trading Using Marginal Abatement Curves.

http://web.mit.edu/globalchange/www/ rpt40.html.

Endres, A. (1997), Negotiating a Climate Convention - The Role of Prices and Quantities. "International Review of Law and Economics", vol. 17, pp. 201-224.

Endres, A. and M. Finus (2002), Quotas May Beat Taxes in a Global Emission Game. "International Journal of Tax and Public Finance" vol. 9, pp. 687-707.

Euractiv (2003), EU News, Policy Positions & EU Actors online. http://www.euractiv.com

Fankhauser, S. (1995), Valuing Climate Change, Earthscan, London.

Finus, M. (2001), Game Theory and International Environmental Cooperation. Edward Elgar, Cheltenham.

Finus, M. (2003a), Stability and Design of International Environmental Agreements: The Case of Transboundary Pollution. Forthcoming in Folmer, H. and T. Tietenberg (eds), International Yearbook of Environmental and Resource Economics, 2003/4, Edward Elgar, Cheltenham, UK.

Finus, M. (2003b), Less Maybe More in International Pollution Control. Preliminary Draft, University of Hagen.

Finus, M. and B. Rundshagen (1998), Toward a Positive Theory of Coalition Formation and Endogenous Instrumental Choice in Global Pollution Control. "Public Choice", vol. 96, pp. 145-186.

Finus, M. and B. Rundshagen (2001), Endogenous Coalition Formation in Global Pollution Control. A Partition Function Approach. Working Paper No. 307, University of Hagen. Revised version forthcoming in Carraro, C. (ed.), Endogenous Formation of Economic Coalitions, Edward Elgar, Cheltenham, UK.

Finus, M. and B. Rundshagen (2003), How the Rules of Coalition Formation Affect Stability of International Environmental Agreements. Preliminary Draft, University of Hagen.

Germain, M. and V. van Steenberghe (2001), Constraining Equitable Allocations of Tradable Greenhouse Gases Emission Quotas by Acceptability. Discussion Paper 2001/5, CORE-Université Catholique de Louvain, Louvain-la-Neuve.

Germain, M., P. L. Toint, H. Tulkens and A. de Zeeuw (2000), Transfers to Sustain Core-Theoretic Cooperation in International Stock Pollutant Control. Revised version of CORE Discussion Paper No. 9832, Center for Operations Research and Econometrics, Université Catholique de Louvain, Louvain.

Hahn, R. W. (1989), A Primer on Environmental Policy Design. Harwood Academic Publishers, Chur et al.

Hillman, A. L. and H. W. Ursprung (1994), Greens, Supergreens, and International Trade Policy: Environmental Concerns and Protectionism. In C. Carraro (ed.), The International Dimension of Environmental Policy, Dordrecht (Kluwer), pp. 75-108.

Hoel, M. (1992), International Environment Conventions: The Case of Uniform Reductions of Emissions. "Environmental and Resource Economics", vol. 2, pp. 141-159.

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

Jeppesen, T. and P. Andersen (1998), Commitment and Fairness in Environmental Games. In: Hanley, N. and H. Folmer (eds.), Game Theory and the Environment. ch. 4, pp. 65-83, Edward Elgar, Cheltenham.

McNutt, P.A. (1996), The Economics of Public Choice, Edward Elgar, Cheltenham.

Michaelowa, A. (1998), Climate Policy and Interest Groups - a Public Choice Analysis. “Intereconomics”, vol. 33, pp. 251-259.

Michaelowa, A. and S. Geiner (1996), Joint Implementation from a Public Choice Perspective. “World Resources Review”, vol. 8, pp. 231-252.

Mueller, D. C. (2003), Public Choice III. Cambridge University Press, Cambridge.

NASA (2002), GDP Inflation Calculator Based on "Budget of the United States Government, Fiscal Year 2001", Historical Tables, Table 10.1, Gross Domestic Product and Deflators Used in the Historical Tables: 1940-2005, http://www.jsc.nasa.gov/bu2/inflateGDP.html.

Nordhaus, W.D. (1994), " Managing the Global Commons ", MIT Press, Cambridge, Ma.

Plambeck, E.L. and C. Hope (1996), An Updated Valuation of the Impacts of Global Warming. In Nakicenovic, N., W. D. Nordhaus, R. Richels, and F. Toth (eds.), Proceeding of the Workshop on Climate Change: Integrating Science, Economics and Policy. IIASA, Laxenburg, Austria, December 1996.

Rubio, S. and A. Ulph (2001), A Simple Dynamic Model of International Environmental Agreements with a Stock Pollutant. Preliminary Version, University of Southampton.

Sandler, T. (1992), Collective Action: Theory and Applications. University of Michigan Press, Ann Arbour.

Schneider, F., and J. Volkert (1999), No Chance for Incentive-Oriented Environmental Policies in Representative Democracies? A Public Choice Analysis. "Ecological Economics", vol. 31, pp. 123-138.

Tol, R.S.J. (1997), A Decision-analytic Treatise of the Enhanced Greenhouse Effect. PhD. Thesis, Vrije Universiteit, Amsterdam.

Tol, R. (2001), Climate Coalitions in an Integrated Assessment Model. "Computational Economics", vol. 18, pp. 159-172.

Weyant, J. (ed.) (1999), The Costs of the Kyoto Protocol: a Multi-model Evaluation. “The Energy Journal”, special issue.

Yandle, B. (1999), Public Choice at the Intersection of Environmental Law and Economics. “European Journal of Law and Economics”, vol. 8, pp. 5-27.

Appendix

Parameter Values

Symbol	Description	Value	Unit	Source
e_{2010}	global emissions in 2010	11.96	Gton CO_2	Nordhaus (1994)
$e_{i,2010}$	regional emissions in year 2010	see Table 1 in section 3	Gton CO_2	own calculation based on Ellerman and Decaux (1998)
d_E	annual growth in global and regional emissions in BAU-scenario	0.153	Gton CO_2	own calculation based on Nordhaus (1994)
$M_{pre\text{-}ind}$	pre-industrial level of CO_2-stock	590	Gton CO_2	Nordhaus (1994)
M_{2010}	stock of CO_2 in 2010	835	Gton CO_2	Nordhaus (1994)
δ	natural annual removal or decay rate of CO_2-stock	0.00866	-	Nordhaus (1994)
ω	airborne fraction of emissions that remain in the atmosphere	0.64		Nordhaus (1994)
r	annual uniform discount rate	0.02	-	assumption
s_i	share of region i in global benefits	see Table 1 in section 3	-	own calculation based on Fankhauser (1995) and Tol (1997)
α_i	abatement cost parameter of region i	see Table 2	-	own calculation based on Ellerman and Decaux (1998)
β_i	abatement cost parameter of region i	see Table 2	-	own calculation based on Ellerman and Decaux (1998)
φ_1	intercept of damage function	-140146	Billion US$	own calculation
φ_2	slope of damage and benefit function	178.331	Billion US$ per Gton	own calculation
γ_D	scale parameter of damage and benefit function	0.027	-	Tol (1997)