Oxford Economic Papers 58 (2006), 209–232 **209**
doi:10.1093/oep/gpi047

The impact of surplus sharing on the stability of international climate agreements

By Hans-Peter Weikard*, Michael Finus† and Juan-Carlos Altamirano-Cabrera‡

*Department of Social Sciences, Wageningen University, Hollandseweg 1 NL-6706 KN Wageningen, The Netherlands; e-mail: Hans-Peter.Weikard@wur.nl
†Department of Economics, University of Hagen, Germany
‡Department of Social Sciences, Environmental Economics and Natural Resources Group, Wageningen University, The Netherlands

This paper analyses stability of coalitions for greenhouse gas abatement under different sharing rules applied to the gains from cooperation. We use a 12-region model to examine internal and external stability of coalitions. We determine and compare stable coalitions under different surplus sharing rules; for example, grandfathering (sharing proportional to current emissions) and a number of equitable rules, i.e. sharing according to historical responsibilities for past emissions. Due to strong free-rider incentives we find only small stable coalitions for all sharing rules examined. We observe that stable coalitions consist of regions with low marginal abatement costs that are attractive partners in any coalition and regions receiving the highest shares of the surplus from cooperation under a particular sharing rule. We find that equitable rules may not be conducive to success: in fact, a grandfathering scheme leads to the most successful coalition in terms of global abatement and global welfare.

JEL classification: D62, D63, Q25.

1. Introduction

The distribution of emission rights ranks high on the international agenda to reach agreements to reduce greenhouse gases (GHGs) and to mitigate climate change. Emission rights in the form of tradable permits are seen as a cost-effective instrument to achieve emission reduction targets for GHGs. However, the introduction of new property rights raises distributional concerns. Grandfathering schemes that allocate tradable permits proportional to emissions in a base year have been criticized for giving advantage to the largest polluters. In the course of the discussion a number of alternative suggestions have been brought forward and are summarized by Rose (1992), Barrett (1992), Kverndokk (1995), and Rose *et al.* (1998). Following Rose *et al.* (1998) it is useful to distinguish three different types of rules for equitable sharing: allocation-based rules which apply to the initial distribution of emission permits, outcome-based rules which apply to the distribution of the net benefits from emission reductions, and process-based rules which comprise criteria for fair decision-making.

This paper deals with outcome-based sharing rules. However, unlike many previous studies on sharing rules such as, for example, Rose *et al.* (1998), we neither simply stipulate the existence of a binding international environmental agreement nor do we ignore free-rider problems. In the absence of an enforcing supra-national body an international environmental agreement has to be self-enforcing (Barrett, 1994). Hence, we study the impact of different sharing rules applied to the gains from cooperation on the success of self-enforcing international climate agreements. Carraro and Siniscalco (1993), Hoel (1992), Barrett (1994), Na and Shin (1998), and others[1] have analysed international environmental agreements as coalition formation games. However, most of this literature uses highly stylized models assuming either identical players or only a very small number of heterogeneous countries. This ignores the conflicting interests of different regions with respect to climate change policies. Therefore, this paper employs a model called STACO (STAbility of COalitions) that combines an empirical module specifying regional benefit and cost estimates of GHG abatement for 12 regions (players) with a game theoretic module that determines stable coalitions (Finus *et al.*, 2006). In empirical models coalition stability has been examined by Botteon and Carraro (1997), Tol (2001), Eyckmans and Tulkens (2003), Eyckmans and Finus (2006), and others. However, the impact of sharing gains from cooperation on coalition stability has not been studied for a larger number of different rules. The work that comes closest to ours are papers by Altamirano-Cabrera and Finus (2006) and Bosello *et al.* (2003). The former also uses the STACO model, but considers allocation-based rules. The latter examines the impact of 'outcome-based equity criteria' on coalition stability in a six-region climate policy model. However, Bosello *et al.* (2003) limit their study to equitable burden sharing, i.e. they focus on the abatement cost side and disregard the distribution of benefits from abatement. Their finding is that equity can be supportive to stability and success of coalitions. By contrast, we arrive at the opposite conclusion and find that different rules of equitable sharing of the gains from cooperation achieve little in terms of global welfare and abatement.

Our analysis of the stability of international climate agreements builds on the cartel formation game with open membership introduced by d'Aspremont *et al.* (1983). The game is a two-stage game. At stage one players decide whether or not to participate in an international agreement. Those who decide to participate form a coalition.[2] We refer to those who do not participate as singletons. At stage two the

[1] See Bloch (2003) for a general survey of coalition formation games and Finus (2003) for a survey focusing on international environmental agreements.

[2] There is at most one coalition ('the cartel') and every player is free to join; this explains the name 'cartel formation game with open membership'. If the cartel consists of a single player, we will call this a 'trivial coalition'.

coalition behaves like a single player; each singleton region and the coalition set emission reduction levels as an optimal response to others' emissions. Payoffs are calculated from costs and benefits of abatement assuming the coalition employs a given sharing rule. The (subgame perfect) equilibria of the game coincide with the set of stable coalitions. A coalition is stable if no member has an incentive to leave the coalition (internal stability) and no singleton player has an incentive to join (external stability). In our specification of the game (see Sections 2 and 4 for details) any coalition of two or more regions will always generate a surplus for its members as compared to the case where all regions are singletons; but it will also generate positive spillovers to non-members. Although there is a surplus, there will still be incentives to free-ride. An important factor determining the free-rider incentives is the distribution of the surplus between coalition members, i.e. the sharing rule applied.

In line with other literature on coalition formation we find only small stable coalitions due to strong free-rider incentives, irrespective of the sharing rule. In general, stable coalitions consist of regions with low marginal abatement costs, which are attractive partners in any coalition, and regions that receive the largest shares of the coalition surplus under a given sharing rule. It turns out that equitable sharing may not be conducive to the success of coalition formation while a grandfathering scheme leads to the most successful coalition in terms of global welfare and abatement. Though we do not claim that the empirical specification of STACO reflects the impacts of climate change in all details, it reflects the main inter-regional differences of GHG abatement and damage costs of climate change. Therefore, our results may be instructive for the future design of climate policies.

The paper is organized as follows. The next section introduces a formal model of coalition formation. Stability of international climate coalitions depends on how the gains from cooperation are shared. We assume that sharing is based on claims and a rule how surplus shares are derived from claims. Section 3 discusses the selection of a surplus sharing rule. We go one step beyond the consideration of the *ad hoc* rules presented by Rose *et al.* (1998) and provide a rationale for the use of proportional sharing. However, an *ad hoc* element remains regarding what constitutes a claim. Section 4 describes the empirical specification of the model introduced in Section 2. Section 5 presents the results of the stability checks for a sample of reasonable claims. Section 6 provides sensitivity analyses. Section 7 concludes.

2. Coalition formation and stability

Coalition formation for GHG abatement is modelled as a two-stage game. Let $N=\{1,2,\ldots,n\}$ be the set of players (regions). At the first stage each player i announces a strategy $\sigma_i \in \{0,1\}$ where $\sigma_i=0$ means that i is not joining the coalition and $\sigma_i=1$ means that i is joining the coalition. A coalition $K \subseteq N$ is a set of k coalition members, $k \equiv |K|$. Since there is at most one non-trivial coalition (i.e. a coalition of at least two members), a coalition structure is sufficiently characterized by coalition K. If $k \leqslant 1$, we call this the singleton coalition structure. If $k = n$ we call this the grand coalition.

The second stage game is a simple transboundary pollution game with two refinements: (i) the coalition behaves like a single player and (ii) there is a redistribution of payoffs within the coalition.

In a simple transboundary pollution game each player i has an initial level of uncontrolled emissions $\bar{e}_i$ and adopts a pollution control strategy (abatement level) $q_i \in [0,\bar{e}_i]$. In the case of GHG abatement q_i is a pure public good. Hence, players receive benefits b_i from total abatement $q = \sum_{i \in N} q_i$ and incur costs c_i of own abatement q_i. Player i's payoff is

$$\pi_i = b_i(q) - c_i(q_i). \tag{1}$$

In the equilibrium each player sets an abatement level q_i^* which is a best response to other players' abatement. At the best response abatement level each player's marginal benefits equal marginal abatement costs. We assume $b_i' \equiv \mathrm{d}b_i/\mathrm{d}q > 0$, $b_i'' \equiv \mathrm{d}^2 b_i/\mathrm{d}q^2 \leqslant 0$, $c_i' \equiv \mathrm{d}c_i/\mathrm{d}q_i > 0$ $c_i'' \equiv \mathrm{d}^2 c_i/\mathrm{d}q_i^2 > 0$, and a regularity property that requires that marginal benefits exceed marginal costs for some small amount of abatement while marginal costs exceed marginal benefits close to the upper bound of the abatement space $\bar{e}_i$. These conditions guarantee a unique interior Nash equilibrium in a public goods game. The Nash equilibrium payoffs are

$$\pi_i^* = b_i(q^*) - c_i(q_i^*). \tag{2}$$

This serves as a benchmark for the following.

The game played at stage two is the simple transboundary pollution game described above where the coalition K and $n-k$ singletons are the players; i.e. there are $n-k+1$ players. Hence, the coalition and each singleton player adopt abatement strategies which are optimal responses to others' emissions. Denote i's abatement level under coalition structure K by q_i^K. For singleton players we obtain the payoffs

$$\pi_i^K = b_i(q^K) - c_i(q_i^K) \quad \text{for all } i \in N \backslash K. \tag{3}$$

For coalition members a sharing rule applies. A sharing rule assigns a share s_i^K of the coalition surplus S^K to every coalition member $i \in K$ such that $\sum_{i \in K} s_i^K = S^K$. The coalition surplus S^K is defined as the joint gain of the coalition members compared with their joint payoff in the benchmark situation of a singletons coalition structure. Formally

$$S^K = \sum_{i \in K} \left(b_i(q^K) - c_i(q_i^K) \right) - \sum_{i \in K} \pi_i^*. \tag{4}$$

The payoff of a coalition member is given by its benchmark payoff plus its share of the coalition surplus:

$$\pi_i^K = \pi_i^* + s_i^K \quad \text{for all } i \in K. \tag{5}$$

The solution concept employed is subgame-perfect equilibrium. The game is solved by backward induction. At stage two of the game optimal responses are characterized by

$$b_i'(q) = c_i'(q_i) \quad \text{for all } i \in N \backslash K \tag{6}$$

and

$$\sum_{i \in K} b_i'(q) = c_i'(q_i) \quad \text{for all } i \in K. \tag{7}$$

The simultaneous solution of the n first order conditions delivers an equilibrium abatement vector that we denote generally by q^{K^*}. In the special cases of $k \leqslant 1$, we denote this vector by q^* and in the case of $k = n$ by q^{N^*}. It is evident that q^* means no cooperation in the context of coalition formation and corresponds to the Nash equilibrium of the simple GHG abatement game without coalition formation. Furthermore, q^{N^*} means full cooperation and corresponds to the global optimum (Mäler, 1989; Folmer and von Mouche, 2000). As q^{K^*} is unique for any coalition K, we obtain second stage equilibrium payoffs from (3) and (5) for any given sharing rule.

It is important to note that the equilibrium abatement vector q^{K^*} is independent of the sharing rule since our coalition formation game assumes transferable utility. Equation (7) implies that the allocation of abatement is efficient within coalition K (Samuelson's condition).[3]

We can now turn to the first stage of coalition formation. An announcement vector σ^* that leads to coalition K is a Nash equilibrium if no coalition member announcing $\sigma_i^* = 1$ is better off by announcing $\sigma_i = 0$, i.e. by leaving coalition K (internal stability), and no singleton announcing $\sigma_i^* = 0$ is better off by announcing $\sigma_i = 1$, i.e. by joining coalition K (external stability). Internal and external stability are formally defined as follows:

Internal stability A coalition K is internally stable if and only if $\pi_i^K \geqslant \pi_i^{K \backslash \{i\}}$, for all $i \in K$.

[3] Since we focus on distributional issues in this paper, we do not discuss mechanisms how an efficient allocation can be achieved. However, we note that it could be achieved with tradable emission permits or an emission tax.

External stability A coalition K is externally stable if and only if $\pi_i^K \geqslant \pi_i^{K\cup\{i\}}$ for all $i \in N\backslash K$.

Hence, the subgame perfect equilibria coincide with the set of stable (i.e. internally and externally stable) coalition structures (Finus *et al.*, 2006).

This completes the discussion of game structure and solution concepts. Before we introduce empirical specifications in Section 4, we turn to sharing rules.

3. Sharing rules

Formally, a surplus sharing problem is a triple $\langle K, \lambda, S\rangle$ where $K \subseteq N$ is a set of k coalition members, $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}_+^k$ is a vector of individual claims of coalition members, and $S \in \mathbb{R}_+$ is the surplus to be shared.[4] Claims are based on characteristics that are considered to be relevant for the sharing problem. This will be discussed below. Let Ω be the set of all surplus sharing problems. A solution to a surplus sharing problem, called sharing rule, is a mapping $\mathcal{R}: \Omega \to \mathbb{R}_+^k$, i.e. a rule $\mathcal{R}$ that assigns a payoff vector $s = (s_1, \ldots, s_k)$ to every surplus sharing problem $\langle K, \lambda, S\rangle$, and $\sum_{i=1}^k s_i = S$. Hence, a sharing rule distributes the entire surplus.[5]

Following Moulin (1987) and, particularly, Pfingsten (1991), we require that a sharing rule satisfies the following properties:

Anonymity For all $i, j \in K$, all $\lambda \in \mathbb{R}_+^k$, and all $S \in \mathbb{R}_+$, $\lambda_i = \lambda_j \Rightarrow s_i = s_j$.

Surplus monotonicity For all $i \in K$, all $\lambda \in \mathbb{R}_+^k$, and all $S, S' \in \mathbb{R}_+$, $S > S' \Rightarrow s_i(K, \lambda, S) \geqslant s_i(K, \lambda, S')$.

Additivity: For all $i \in K$, all $\lambda \in \mathbb{R}_+^k$ and all $S, S_0, S_1 \in \mathbb{R}_+$ such that $S = S_0 + S_1$, $s_i(K, \lambda, S) = s_i(K, \lambda, S_0) + s_i(K, \lambda, S_1)$.

Separability: For all $i \in K$, all $H \subset K$ and $H \neq \varnothing$, all $\lambda, \lambda' \in \mathbb{R}_+^k$, and all $S, S' \in \mathbb{R}_+$, $\lambda_i = \lambda'_i$ for all $i \in H$ and $\sum_{i\in H} s_i(K, \lambda, S) = \sum_{i \in H} s_i(K, \lambda', S') \Rightarrow s_i(K, \lambda, S) = s_i(K, \lambda', S')$.

Anonymity requires equal treatment of equals. Surplus monotonicity says that no coalition member should loose if the surplus increases. Additivity says that payoffs should not change if the surplus is distributed in two instalments instead of one. Separability is a subgroup consistency requirement which says that individual payoffs in every subgroup depend only on the claims of the players in the

[4] We drop the superscript K on S in this section.

[5] This is a formal definition of a sharing rule when claims are given; in other sections of the paper, we use the term 'sharing rule' in a broader sense reflecting also claims.

subgroup and the subgroup's surplus. Anonymity and Surplus monotonicity are hardly debatable. We also argue that Additivity applies to the case at hand. As the true damages of climate change, and, hence, the true benefits from global abatement become known at a later stage, the distribution should not depend upon the pattern of how benefits become available. The case for Separability is that it should not matter for the final outcome whether a player receives its share of the coalition surplus directly or as a share of the payoff of a subcoalition.

Pfingsten (1991) has shown that these properties characterize a family of sharing rules.[6] He has shown that a sharing rule $\mathcal{R}$ satisfies Anonymity, Surplus monotonicity, Additivity, and Separability if and only if $\mathcal{R}$ is either (i) equal sharing: $s_i(K, \lambda, S) = (1/k)S$, (ii) proportional sharing $s_i(K, \lambda, S) = (\lambda_i / \sum_{j \in K} \lambda_j)S$, or (iii) a combination of (i) and (ii) $s_i(K, \lambda, S) = \alpha(1/k)S + (1-\alpha)(\lambda_i / \sum_{j \in K} \lambda_j)S$, where $0 < \alpha < 1$.

As Anonymity, Surplus monotonicity, Additivity and Separability are convincing properties in the context of GHG abatement coalitions, this result characterizes the set of reasonable sharing rules.

In what follows, we consider a set of eight sharing problems which differ with respect to what constitutes a claim. Although it is possible to apply different rules to these sharing problems (equal sharing, proportional sharing and combinations), we focus on proportional sharing.

1. Egalitarian claims: $\lambda_i = \lambda_j$, for all i, j

All players (regions) have equal claims. Egalitarian claims reflect the principle 'one region, one vote'. This does not seem to be convincing in the case of climate coalitions of heterogeneous regions. However, we include this case as a benchmark because proportional sharing under egalitarian claims coincides with equal sharing.

2. Income claims: $\lambda_i = GDP_i$

where GDP_i is region i's gross national product in a base year. This rule has been also dubbed 'horizontal equity' by Rose *et al.* (1998). One appealing feature of this rule is that it maintains relative welfare positions.

3. Population claims: $\lambda_i = pop_i$

where, pop_i is region i's population in a base year. The motivation for this rule is that individuals should have equal rights to the global commons and, hence, that gains from cooperation should be distributed evenly across the global population.

[6] In his proof, Pfingsten (1991) also uses a property called 'No advantageous reallocation' which requires that the coalition surplus is independent of the distribution of claims. Obviously, this always holds in the GHG abatement game analysed in this paper.

4. Ability-to-pay claims: $\lambda_i = (GDP_i/pop_i)^{-\gamma}$, $\gamma > 0$

Regions with a lower per capita income have larger claims. Under this rule climate policy is used as a vehicle to reduce income inequality. Thus, the motivation of this rule stems from outside climate policy and is guided by some principle of 'international justice'.

5. Emission claims: $\lambda_i = e_i$

where, e_i are region i's emissions in a base year. Emissions claims can be interpreted as historical emission rights.

6. Inverse emission claims: $\lambda_i = e_i^{-\gamma}$ with $\gamma > 0$

Regions with higher emissions receive a lower share of the gains from cooperation. These claims reflect historical responsibilities for the accumulated stock of GHGs.

7. Damage cost claims: $\lambda_i = d_i$

where, d_i is the net present value of region i's damages from climate change. After implementation of abatement policies, some climate change damages will still remain. Those who suffer larger damages should receive a larger compensation.

8. Abatement cost claims: $\lambda_i = c_i$

where, c_i is the net present value of region i's abatement costs. The coalition surplus is interpreted as a return to investments in abatement. Those who bear larger costs should be entitled to a larger share of the coalition surplus.

Of course, a longer list of possible claims could be generated. In addition to egalitarian claims that serve as a benchmark we include income, population, and ability-to-pay claims because they have been extensively discussed by Rose *et al.* (1998). Emissions claims are probably the most prominent claims. They are the outcome-based analogue to a grandfathering scheme in the context of emission permits. Inverse emission claims, which reflect historical responsibilities, are less prominent in economic analysis, but have received some attention in philosophy (Gosseries, 2003; Weikard, 2004). We have included damage costs and abatement cost claims because they reflect different views on compensation. Marginal damage cost claims seem worth considering as they have been discussed in the literature (cf. Chander and Tulkens, 1995). However, our empirical results are derived from a linear damage cost function. In this case, the use of marginal damage cost claims leads to the same result as the use of damage cost claims. Marginal abatement cost claims lead to equal sharing because the optimal abatement strategy for the coalition requires equal marginal abatement cost for all coalition members.

4. Empirical model and data

In order to examine the impact of sharing rules on coalition stability, we adopt the STAbility of COalitions model, STACO, introduced by Finus *et al.* (2006). STACO is designed for the game theoretical analysis of coalition stability in the context of international climate agreements. Since free rider problems are particularly pronounced in the context of many players, STACO comprises 12 world regions. The following regions are considered: United States (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 all remaining other countries (ROW). STACO considers a time horizon of 100 years and adopts the simplest possible benefit and cost functions that capture the main features of the GHG problem. STACO assumes linearly growing baseline emissions (regional emission shares are reported in Table 2, column 5), stationary abatement strategies, and constant abatement costs. Parameters are calibrated such that the relevant magnitudes of emissions and concentrations of greenhouse gases are well reflected. The STACO model uses a linear approximation of the damage cost function of the DICE model introduced by Nordhaus (1997). Moreover, the damage cost function is rescaled using estimates of Tol (1997). Global benefits from abatement are defined as avoided damages. Regional benefits are calculated as shares of global benefits from abatement based on estimates from Fankhauser (1995) and Tol (1997). The shares are reported in Table 2, column 7.[7] Because STACO uses a linear benefits function, marginal benefits are constant and are reported in Table 1, column 4. Recall that for each region marginal benefits equal marginal abatement costs for the singletons coalition structure. The payoff function for region i is given by

$$\pi_i = \mu_i \delta_b \gamma \sum_{i=1}^{12} q_i - \delta_c(\xi_i q_i^3 + \zeta_i q_i^2) \tag{8}$$

where μ_i ($\sum_{i=1}^{12} \mu_i = 1$) is the regional benefits share, δ_b captures discounting of benefits from abatement and the effect of decay of GHGs, γ is a scaling parameter for global benefits, and δ_c captures discounting of abatement costs; the abatement cost function $\xi_i q_i^3 + \zeta_i q_i^2$ and the regional parameters $\xi_i, \zeta_i > 0$ are taken from Ellerman and Decaux (1998). In our base case we assume a discount rate of 2%. Impacts of a change of the discount rate are examined in a sensitivity analysis in Section 6. Further details concerning calibration of parameters are provided by Finus *et al.* (2006).

[7] Notice that when damage costs are linear and benefits are defined as avoided damages, benefits shares equal damage cost shares.

Table 1 Benchmark cases: singleton and grand coalition structure

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Singletons coalition structure				Grand coalition structure		
Region	Marginal abatement costs at 50 Mton/yr ($/ton)	Emissions reduction (Mton/ year)	Marginal abatement costs ($/ton)	Total abatement costs over 100 years ($bn)	Emissions reduction (Mton/ year)	Marginal abatement costs ($/ton)	Total abatement costs over 100 years ($bn)
USA	1.40	162.3	8.46	53.33	379	37.4	513
JPN	55.84	7.7	6.45	2.44	36	37.4	63
EEC	5.82	66.2	8.83	24.22	161	37.4	229
OOE	8.94	19.0	1.29	0.82	102	37.4	127
EET	9.56	9.3	0.49	0.18	102	37.4	130
FSU	2.57	49.6	2.52	4.24	193	37.4	242
EEX	9.98	7.9	1.12	0.43	124	37.4	188
CHN	0.59	154.9	2.32	16.09	956	37.4	1,348
IND	3.31	33.6	1.87	2.73	216	37.4	295
DAE	13.20	5.4	0.93	0.24	102	37.4	155
BRA	787.81	0.2	0.57	0.00	7	37.4	12
ROW	4.00	37.2	2.54	3.95	185	37.4	250
World		553.2		108.68	2,563	37.4	3,553

Source: Finus *et al.* (2005), own calculations.

Table 2 Overview of claims

(1) Regions	(2) Income[(a)(b)] (US$bn)	(3) Population[(b)(c)] (million inhabitants)	(4) Ability-to-pay[(d)] $(US\$)^{-1}$	(5) Emissions[(e)] (Gton)	(6) Inverse emissions[(f)] $(Gton)^{-1}$	(7) Damage costs[(g)] (%)
USA	8,845	305	431	2.42	0.41	22.6
JPN	5,584	124	386	0.56	1.80	17.3
EEC	9,579	375	445	1.40	0.71	23.6
OOE	1,902	142	523	0.62	1.61	3.5
EET	405	120	736	0.51	1.93	1.3
FSU	501	287	863	1.00	1.00	6.7
EEX	1,650	1,602	1,000	1.22	0.82	3.0
CHN	1,021	1,340	1,057	2.36	0.42	6.2
IND	458	1,145	1,257	0.63	1.56	5.0
DAE	972	207	679	0.41	2.47	2.5
BRA	774	190	703	0.13	7.81	1.5
ROW	1,119	584	852	0.70	1.43	6.8
WORLD	32,810	6,421	–	11.96	–	100.0

Notes: (a) Data refer to the level of GDP in 2010 expressed in 1985 US$. The global figure for 2010 is taken from DICE and regional shares are computed from table 1.1 of World Bank (2002). (b) Data for individual countries was aggregated into our 12 regions following Babiker *et al.* (2001). (c) Data refers to the level of population in 2010 that has been extrapolated from 2000 using table 2.1 of World Bank (2002). (d) Computed from columns 2 and 3 for $\gamma = 0.25$. (e) Own calculations from STACO; emissions in 2010. (f) Computed from column 5 for $\gamma = 1$. (g) STACO calibration, see Finus *et al.* (2006).

Table 1 summarizes some basic settings of the STACO model. Column 2 reports discounted marginal abatement costs for a uniform abatement level across regions. Furthermore, Table 1 reports emissions reductions for the singletons coalition structure (column 3) and the corresponding marginal and total abatement costs (columns 4 and 5). Emissions reductions and marginal and total abatement costs are also reported for the grand coalition (columns 6–8). It can be seen from column 2 that CHN has the lowest marginal abatement costs followed by USA and FSU, whereas BRA has by far the highest for a uniform abatement level. CHN, USA and FSU have high emissions levels (see Table 2, column 5) and cheap abatement options, while BRA's abatement options are expensive due to low emissions levels. For the singleton coalition structure the picture changes. EET and BRA have the lowest marginal abatement cost while EEC and USA have the highest. In this case each region equates marginal abatement costs with marginal damage costs which causes USA and EEC to adopt high levels of abatement while BRA chooses to abate very little. In the grand coalition 37% of global abatement will be undertaken by CHN since it provides the cheapest abatement options. Therefore, it can be conjectured that CHN will be an attractive partner in any climate coalition. Because in the grand coalition each region equates marginal abatement costs to the sum of marginal benefits of all regions, all regions face the same marginal abatement costs. Note that discounted marginal abatement costs of \$37 per ton CO_2 in the global optimum (column 7) is a figure much in line with results obtained from more sophisticated climate models as, for instance, Plambeck and Hope (1996).

Individual payoffs for coalition members follow from the particular coalition surplus sharing scheme applied. Individual shares are proportional to claims. Table 2 presents the input data for the claims specified in Section 3. The table does not report egalitarian claims which are the same for all. Also the table does not report abatement cost claims which are sensitive to coalition structure.

Coalition membership is most attractive for a region if it has a high claim and, hence, receives a large share of the surplus. So we expect to find EEC and USA in a stable coalition if surplus sharing is according to income or damages. EEX and CHN receive the largest shares under population claims, CHN and IND under ability-to-pay claims, USA and CHN under emissions claims, and BRA and DAE under inverse emission claims.

5. Results and discussion

We start the discussion of results by noting three general features of our model.

First, note that a trivial coalition where only a single player announces $\sigma_i = 1$ $(k = 1)$ is not effective. The singletons coalition structure emerges. Therefore, announcements $\sigma_i = 0$ for all $i \in N$ constitute a Nash equilibrium as single deviations will never improve the payoff. Hence, the singletons coalition structure is stable.

Second, coalition formation in the GHG abatement game specified in Section 2 generates positive spillovers to non-members of K. The coalition applies

Samuelson's condition (eq. (7)). Hence it abates more than its members abate under a singletons coalition structure. Singletons either maintain (no-leakage) or reduce (leakage) their abatement effort; see Finus and Rundshagen (2003) for a proof. Thus, if a coalition forms, singletons enjoy higher benefits from global abatement while they have the same or reduced abatement costs.[8]

Third, under the STACO specification benefits are linear in abatement. Under linear benefits (i.e. constant marginal benefits) players have dominant abatement strategies.[9] No coalition or singleton region will adjust its abatement if others change theirs because there is no change in marginal benefits; see conditions (6) and (7). The fact that regions have dominant strategies in the GHG abatement game implies that there is no leakage. Members of a non-trivial coalition will abate more compared with the singletons coalition structure. This additional abatement is not offset by less abatement of the remaining singletons, as they have dominant strategies. We explore the case of leakage in Section 6.

The STACO model is used to generate the payoffs for every possible coalition structure ($2^{12} - 12 = 4084$ in a 12 regions model) for the sharing schemes described above. STACO performs a stability check and identifies internally stable, externally stable, and stable coalitions. The results for the eight sharing schemes and the benchmark cases (singleton coalition structure and grand coalition) are summarized in Table 3. The stable coalitions for each scheme are listed in column 2. Column 3 reports global annual emission reduction and columns 4–6 display costs, benefits and net benefits from abatement. Note that a considerable amount of net benefits is already obtained under the singleton coalition structure. The additional net benefits due to coalition formation are reported in column 7 as the sum of coalition surplus and external benefits (spillovers). This is expressed in relative terms in column 8.

There are several findings. The first observation from Table 3 is that there are non-trivial stable coalitions for all 8 sharing schemes. For comparison, Finus *et al.* (2006) do not find any stable coalition for the STACO model in the absence transfers. The design of the transfer schemes is important. Our result is also strikingly different from the findings of Altamirano-Cabrera and Finus (2006) who consider transfers under allocation based sharing. In their setting sharing schemes are applied to emission permits and not, as in this paper, to the net benefits from coalition formation. Altamirano-Cabrera and Finus (2006) find a total of four stable coalitions for grandfathering schemes and they do not find any stability for any of the 'equitable rules' they consider. To understand the difference, notice that the following holds.

[8] Note that the spillover effect also implies an implicit punishment for coalition members considering to leave the coalition. Leaving does not only lower abatement costs but also lower benefits due to lower global abatement. In this sense our coalition formation model mimics a trigger strategy in a repeated game. After player i has left coalition K, the remaining coalition members reoptimize their strategies and play their best response.

[9] Cf. Folmer and von Mouche (2000), Proposition 2.

Table 3 Overview of results

(1) Sharing schemes (benchmark cases)	(2) Coalition members	(3) Global annual emissions reduction (Mton)	(4) Global abatement costs ($USbn over 100 years)	(5) Global benefits ($USbn over 100 years)	(6) Global net benefits ($USbn over 100 years)	(7) Coalition surplus + external net benefits ($USbn over 100 years)	(8) Global net benefits relative to grand coalition (%)
(only singletons)		553	109	2,069	1,960	0 + 0	0.0
(Grand coalition)*	(all regions*)	2563	3,553	9,584	6,031	4071 + 0	100.0
Egalitarian	EET, CHN, IND	711	159	2,658	2,499	22 + 516	13.2
Income	EEC, CHN	870	311	3,253	2,942	151 + 831	24.1
Population	EEX, CHN	620	127	2,317	2,190	4 + 226	5.7
Ability-to-pay	EET, FSU, CHN	731	172	2,735	2,563	32 + 571	14.8
	EET, EEX, CHN	665	140	2,485	2,346	12 + 374	9.5
	EET, CHN, IND	711	159	2,658	2,499	22 + 516	13.2
Emissions	USA, EET, EEX, CHN	1030	436	3,854	3,418	264 + 1194	35.8
Inverse emissions	EET, BRA	559	109	2,090	1,981	0.2 + 21	0.5
	CHN, BRA	582	116	2,176	2,059	1 + 98	2.4
Damage costs	USA, CHN	874	314	3,270	2,956	142 + 854	24.5
	EEC, CHN	870	311	3,253	2,942	151 + 831	24.1
Abatement costs	USA, CHN	874	314	3,270	2,956	142 + 854	24.5
	JPN, CHN	796	237	2,976	2,739	85 + 694	19.1
	OOE, CHN	626	129	2,341	2,212	6 + 246	6.2
	FSU, CHN	683	154	2,553	2,398	17 + 421	10.8
	EEX, CHN	620	127	2,317	2,190	4 + 226	5.7
	CHN, IND	662	143	2,477	2,334	11 + 363	9.2
	CHN, ROW	683	155	2,555	2,400	17 + 423	10.8

*The grand coalition is not stable irrespective of the sharing scheme.

Proposition 1 In a GHG abatement game with coalition formation and proportional surplus sharing if benefits from abatement are linear and if claims are non-negative, all non-trivial coalitions K are profitable, i.e. for all $\pi_i^{K^*} \geqslant \pi_i^*$ for all $i \in K$.

Proof GHG abatement is a global public good. Consider coalitions K' and K with $K' \subset K$ and $K' \neq \varnothing$. It holds that $q^{K^*} > q^{K'^*}$ as the larger coalition will abate more (Samuelson's condition) and the singletons maintain their dominant strategy. Hence, it also holds that the coalition surplus is increasing in coalition size, $S^{K^*} > S^{K'^*}$. Hence, there is always a strictly positive surplus for every coalition K with $k > 1$. With non-negative claims and proportional sharing it holds that $s_i^K \geqslant 0$. Hence, $\pi_i^{K^*} = \pi_i^* + s_i^K \geqslant \pi_i^*$. □

Proposition 2 In a GHG abatement game with coalition formation and proportional surplus sharing if benefits from abatement are linear and if claims are non-negative, all two-player coalitions are internally stable.

Proof Consider coalitions K' and K with $K' \subset K$, $k' = 1$, and $k = 2$. Internal stability follows immediately from Proposition 1. □

A corollary of Proposition 2 is that for coalitions K' and K with $K' \subset K$, global welfare is higher in K than in K'. This follows from the facts that (i) the coalition surplus in K is higher than in K' and that (ii) all players that do not belong to K are better off under K than under K' due to positive spillovers. Thus, coalition formation generates a global welfare gain. It also follows that the grand coalition maximizes global welfare.

Outcome-based rules guarantee profitability. Every coalition member receives at least the payoff obtained under the singletons coalition structure. Moreover, every two-player coalition is at least internally stable. This is generally not the case for allocation based-rules considered by Altamirano-Cabrera and Finus (2006) and it is also not the case in the setting of Bosello *et al.* (2003).

The second observation is that singletons receive large external net benefits for all coalitions listed (see Table 3, column 7). This indicates large positive spillover effects and, hence, strong free-rider incentives. This explains why we find only small stable coalitions that comprise only two regions under most sharing schemes and that the largest stable coalition comprises four regions. Nevertheless, the sharing schemes examined here enhance stability and the success of cooperation.

Third, egalitarian claims, population claims, ability-to-pay claims and inverse emission claims that reflect some notion of equitability are not very successful in terms of global emission reduction and in terms of global net benefits. Abatement cost claims give a mixed picture. Sharing according to income claims and damage cost claims are more successful. The best result is obtained for emission claims. The stable coalition comprises USA, EET, EEX, and CHN and achieves about 36% of the gains that the grand coalition would achieve. This indicates that equitable sharing may not be helpful to achieve successful international climate agreements in the presence of strong free-rider incentives. The reason is that most of the equitable rules allocate large shares of the gains from cooperation to developing countries. This gives too little incentives to industrialized regions to participate in an international climate agreement.

Fourth, coalition membership seems to follow some general pattern. We observe that CHN is a member in every stable coalition except for a case with inverse emission claims. CHN's low abatement cost options generate a high coalition surplus of which it receives a sufficient share under almost every rule. Therefore, (almost) every two-player coalition not involving CHN is externally unstable. Hence, if a two-player coalition is stable, it is likely to involve CHN. From Proposition 2, we know that every two-player coalition is internally stable. However, generally, it is also attractive for other regions to join a coalition including CHN. This is particularly true for regions with large claims. Thus, if we find stable two-player coalitions, they will consist of CHN and the region with the highest claims. This pattern applies in a straightforward manner to income claims ({EEC, CHN}), population claims ({EEX, CHN}), inverse emission claims ({CHN, BRA}), and damage cost claims ({USA, CHN} and {EEC, CHN}). For equal sharing, USA or EEC are not part of a stable coalition with CHN since CHN would receive a too large share of the surplus.

The case of inverse emissions claims involves stable coalitions with BRA. Here BRA has by far the largest claim. This makes it attractive for BRA to join any existing coalition. Hence, coalitions not involving BRA will be externally unstable. However, most coalitions with BRA are unattractive for other coalition partners and will be internally unstable, unless the coalition is of size 2 (see Proposition 2). The two stable two-player coalitions we find in this case achieve very little because BRA faces high marginal abatement costs.

In the remainder of this section we examine the factors relevant to the composition of stable coalitions in more detail. We seek to identify regions that are more likely to join a coalition than others. In our GHG abatement game incentives of regions to cooperate depend on three factors: on marginal abatement costs, on marginal benefits, and on claims to the coalition surplus. For the decision whether or not to join a coalition a region compares its share of the surplus when joining a coalition with its free-rider surplus. First, consider the impact of marginal abatement costs. Regions which have low marginal abatement costs contribute more to the coalition surplus since they provide cheap abatement options. Hence, other things equal, we would expect to find the regions with the lowest marginal abatement costs in a coalition. Second, the impact of marginal benefits is ambiguous. On the one hand, high marginal benefits stimulate coalition partners to abate more which contributes to a higher coalition surplus. On the other hand high marginal benefits are an incentive to free-ride. One can define a free-rider surplus as the product of marginal benefits from abatement and the additional abatement of the coalition (compared to the singletons coalition structure).[10] We presume that high

[10] Marginal benefits are assumed to be constant and are given in Table 1, column 4. Note that marginal benefits equal marginal abatement costs in the singletons coalition structure.

marginal benefits cause stronger incentives to free-ride than incentives to join the coalition. This is because the additional surplus of joining will have to be shared with other coalition members. Other things being equal a region is more likely to join a coalition if its marginal benefits are low. Third, with unequal claims, a region is more likely to join a coalition if its claims are high.

We use this argument to construct a rough indicator for the relative advantage from coalition membership. We use the following ingredients: (i) marginal abatement cost (at 50 Mton per year, Table 1, column 2), (ii) marginal benefits (Table 1, column 4), and (iii) the share of total claims (Table 2). Rescaling the cost and benefits parameters, we propose the following coalition membership index I:

$$I = \frac{\ln(1+b_i')}{\ln(1+c_i')} \cdot \frac{1}{(\ln(1+b_i'))^2} \cdot \frac{\lambda_i}{\sum_{j \in N} \lambda_j}.$$

The first factor captures surplus size. It is higher if a region's marginal abatement costs are low and if its marginal benefits are high. The second factor captures (inverse) free-rider incentives which are lower if a region's marginal benefits are high. Marginal benefits receive a larger weight in the free-rider effect as compared to the surplus size effect as argued above. The third factor captures the size of a region's share in a coalition. In summary, a region is more likely to be a coalition member if it has a high coalition membership index, i.e. if its marginal abatement costs are low, if its marginal benefits are low, and if its share of the surplus is high. Of course, such indicator cannot work precisely as marginal abatement cost and the share of a region's surplus will be coalition dependent. A general coalition membership index cannot be constructed as this requires to attach weights to each component of the index which will differ between claim types. However, based on parameters c_i', b_i', and λ_i we can obtain a partial ordering of coalition membership: If region i is a coalition member, then region j with $c_j' < c_i'$, $b_j' < b_i'$ and $\lambda_i > \lambda_i$ will also be a coalition member. If region i is not a coalition member, then region j with $c_j' > c_i'$, $b_j' > b_i'$ and $\lambda_j < \lambda_i$ cannot be a coalition member either.

The index we suggest is reported in Table 4. For equal sharing the highest coalition membership indices are reported for CHN, EET, and IND. These regions form the only externally stable coalition of the about 100 internally stable coalitions for the case of equal sharing. This confirms our expectation. In the case of income claims USA has a higher index than CHN. In this case the index identifies only EEC correctly as a coalition member. For the cases of population claims, ability-to-pay claims, inverse emission claims and damage cost claims the index performs well, identifying correctly members of stable coalitions. For emission claims three of the four coalition members are correctly identified.

With emissions claims we find the most successful stable coalition ({USA, EET, EEX, CHN}). As can be seen from Table 5 the success of the coalition for the global surplus depends largely on the presence of both, USA and CHN, in the coalition. The three-players coalitions {EET, EEX, CHN} and {USA, EET, EEX} are less

Table 4 Coalition membership index*

(1) Regions	(2) Equal sharing	(3) Income	(4) Population	(5) Ability-to-pay	(6) Emissions	(7) Inverse emissions	(8) Damage cost
USA	0.22	*0.73*	0.13	0.13	**0.54**	0.05	**0.61**
JPN	0.05	0.11	0.01	0.03	0.03	0.05	0.11
EEC	0.10	**0.35**	0.07	0.06	0.14	0.04	0.29
OOE	0.23	0.16	0.06	0.16	0.14	0.20	0.10
EET	**0.47**	0.07	0.11	**0.46**	***0.24***	**0.49**	0.07
FSU	0.28	0.05	0.15	0.32	*0.28*	0.15	0.22
EEX	0.25	0.15	**0.73**	0.33	**0.30**	0.11	0.09
CHN	**0.79**	***0.30***	**1.99**	**1.13**	**1.88**	0.18	**0.59**
IND	**0.29**	0.05	0.61	**0.48**	0.18	0.25	0.17
DAE	0.25	0.09	0.10	0.23	0.10	0.34	0.08
BRA	0.15	0.04	0.05	0.14	0.02	**0.63**	0.03
ROW	0.22	0.09	0.24	0.25	0.15	0.17	0.18

*Bold figures indicate members of stable coalitions; italic figures indicate two cases where membership is not correctly predicted by the index.

Table 5 Results for coalition structure {USA, EET, EEX, CHN} and neighbouring coalitions under emissions claims

Regions	EET, EEX, CHN	USA, EEX, CHN	USA, EET, CHN	USA, EET, EEX	USA, EET, EEX, CHN
	share of coalition surplus (bold) or free-rider surplus ($USbn over 100 years)				
USA	94	**83**	**85**	**26**	**96**
JPN	72	265	248	66	308
EEC	98	362	339	90	421
OOE	14	53	50	13	62
EET	**2**	20	**19**	**6**	**22**
FSU	28	104	97	26	120
EEX	**4**	**45**	43	**14**	**52**
CHN	**7**	**81**	**83**	24	**94**
IND	21	77	72	19	89
DAE	10	38	36	10	44
BRA	6	23	22	6	27
ROW	28	104	98	26	121
World	386	1,256	1,191	326	1,458

successful than {USA, CHN}, which achieves a global surplus of US$996bn over 100 years (not reported in the table). In the case of emissions claims CHN has strong incentives to join as it has high claims (and obtains a large share of the surplus when joining) and it has low marginal abatement costs. In fact, no coalition that does not include CHN is externally stable. USA joins CHN because under emission claims USA receives the largest share. EEC has the third largest claim, but has a strong free-rider incentive. USA and CHN are joined by EET and EEX who receive lower shares than EEC but have less incentive to free-ride. The simple intuitive explanation why emission based claims are more successful than any alternative rule considered here is as follows. A high level of emissions is linked to better opportunities for abatement and, hence, low marginal abatement costs. Coalitions that include regions with lower marginal abatement costs create a larger surplus. Under emissions claims these regions are encouraged to join a coalition.

6. Sensitivity analysis

In order to check robustness of results we have conducted sensitivity analyses with respect to parameters and with respect to the functional form of the benefits. The objective of our sensitivity analysis is to check the robustness of our qualitative results and conclusions, rather than to confirm the stability of particular coalitions. The analysis of parameter sensitivity has been restricted to global parameters: the global level of benefits and abatement costs as well as the discount rate. If the assumption of a linear benefit function is relaxed, regions do no longer have dominant abatement strategies. A region's abatement efforts will partly be offset by less abatement in other regions. There is leakage. We turn to parameter sensitivity first.

Sensitivity of parameters Any empirical analysis must acknowledge uncertainty of parameter values. In models with many parameters it is impossible to explore all possibilities. Therefore, we restrict the analysis to global parameters in our model. We analyse a uniform change of benefits, abatement costs, and the discount rate. The analysis is simplified by the fact that a change of the discount rate has the same implications as a uniform change of the benefits and cost parameters. To see this, recall that the payoff function in STACO is given by $\pi_i = \mu_i \delta_b \gamma \sum_{i=1}^{12} q_i - \delta_c(\xi_i q_i^3 + \zeta_i q_i^2)$. First, consider abatement costs. A higher (lower) discount rate just scales down (up) the abatement costs because we assume a stationary abatement path and, hence, constant abatement costs. The parameter δ_c is the accumulated discount factor and works like a scaling parameter. Next, consider benefits. A change of the discount rate has a similar scaling effect captured by the parameter. However, due to stock effects, most of the benefits of abatement are obtained in the future. Therefore, a change of the discount rate has a stronger scaling effect on benefits than on abatement costs. Hence, an increase of the discount rate has the same effect as lowering benefits and/or increasing abatement costs. Lowering the discount rate has the same effect as increasing benefits and/or lowering abatement costs. This allows us to focus our sensitivity analysis on the discount rate.[11]

[11] If there were no stock effect, all benefits would stem from current abatement. In that case a change of the discount rate would not affect the optimality conditions, as equally strong scaling effects prevail on the both sides, the benefits and the cost side.

In our base case we assume a discount rate of 2%. In our sensitivity analysis we apply discount rates of 1%, 3%, and 5%. As explained above, a 1% discount rate can be interpreted as a rise in benefits and abatement costs as compared to the base case. Due to the stock effect, benefits rise more than abatement costs. The aggregate effect is the same as scaling up benefits by a factor 1.18. For 3% and 5% discount rates the aggregate effect is the same as scaling benefits down by a factor 0.85 and 0.62, respectively. A summary of our findings follows.

For discount rates of 1% and 3% we find the same stable coalitions for all eight sharing rules. If the discount rate is raised to 5%, three changes occur. First, we find an additional stable coalition consisting of CHN and IND for population claims. Second, the coalition of EET, EEX, and CHN is no longer (internally) stable for ability-to-pay claims. Third, the coalition EEX and CHN is no longer (externally) stable for abatement cost claims. In this case USA wants to join EEX and CHN. However, this enlargement is not internally stable as EEX would leave. This leads to a stable coalition of USA and CHN. This shows that our analysis is relatively robust with respect to a change of global parameters. Our general observations are maintained. Only small coalitions are stable, equitable sharing schemes are not conducive to successful international climate agreements, and grandfathering (emissions claims) performs best.

Also the general pattern of membership and its explanation are still valid. The region with the lowest marginal abatement costs forms a coalition with those regions that receive the highest shares under the respective sharing scheme. Furthermore, the coalition membership index performs equally well as in the base case.

Non-linear benefit functions As pointed out in Section 5, linear benefit functions imply dominant strategies. Hence, if a non-trivial coalition $K \subset N$ forms and increases its aggregate abatement efforts compared to non-cooperative levels, the additional abatement is not offset by a reduction of abatement efforts of the singletons. There are no leakage-effects. This may be regarded as the most favourable conditions for coalition formation (Carraro, 2000). Therefore, it is worthwhile to consider the effect of non-linear benefit functions. Based on standard results for transboundary pollution games without coalition formation we expect that this implies leakage also for the case of coalition formation. The following result shows that this is indeed the case.

Proposition 3 In the GHG abatement game with coalition formation, if benefits functions of global abatement are concave and if individual abatement cost functions are strictly convex, then there is (i) no leakage if benefit functions are linear and (ii) leakage if benefit functions are strictly concave, although leakage is never complete.

Proof Consider coalition $K \subset N$ with aggregate abatement q_K. Denote the aggregate abatement of the singletons ($N \backslash K$) by q_{-K}. Then, total differentiation of the first order conditions in eq. (7) (which are the implicit best response functions of coalition members) and appropriate manipulation gives the slope of the aggregate best response function of coalition K:

$$\frac{\mathrm{d}q_K}{\mathrm{d}q_{-K}} = -\frac{b''_K\left(1+\sum_{j=2}^{k}(c''_i/c''_j)\right)}{b''_K\left(1+\sum_{j=2}^{k}(c''_i/c''_j)\right)-c''_i}, \tag{9}$$

where $i, j \in K$, c''_i denotes the second order derivative of i's abatement cost function, and b''_K denotes the sum of the second order derivatives of the benefit functions of all coalition members and the arguments in these functions have been omitted. For linear benefit functions $b''_K = 0$ holds and hence $(\mathrm{d}q_K/\mathrm{d}q_{-K}) = 0$. For strictly concave benefit functions $b''_K < 0$ holds. Since $c''_i > 0$, it follows that $-1 < (\mathrm{d}q_K/\mathrm{d}q_{-K}) < 0$. □

Proposition 3 shows that there is an upper and a lower bound on leakage effects. At the lower bound if benefit functions are linear, no leakage occurs. There is leakage in all other cases. However, leakage has an upper bound. Additional abatement efforts are never completely offset as $-1 < (\mathrm{d}q_K/\mathrm{d}q_{-K})$. Note that for a trivial coalition where $k = 1$ eq. (9) collapses to $(\mathrm{d}q_i/\mathrm{d}q_{-i}) = -(b''_i/(b''_i - c''_i))$.

The fact that leakage is never complete has important implications for coalition formation. With complete leakage no non-trivial coalition can be stable. Coalition members would have higher abatement costs but the same (or lower) benefits than in the singleton coalition structure. Hence, a coalition cannot be internally stable. Note that complete leakage may occur in transboundary pollution games when pollution is not uniformly distributed as for instance in Mäler's (1989) acid rain game.

By varying the slope of the marginal benefit functions different degrees of leakage can be captured. Regardless of the specific functional form of a downward sloping marginal benefit function, it will either lie above (in some range) or below (in some range) the horizontal marginal benefit function of the base case. Thus, changing the horizontal marginal benefits curve into downward sloping one will not only capture leakage effects, as intended, but will also have level effects as marginal benefits will necessarily be below or above their base case level.

With this caveat in mind we consider the benefit function $b_i = \mu_i \delta_b \gamma \sum_{j=1}^{12} q_j^{\alpha}$ with $0 < \alpha < 1$. This implies a strictly convex marginal benefit functions. For $\alpha = 1$ we obtain the base case.

We consider $\alpha = 0.99$ and $\alpha = 0.95$ and we recalibrate the parameter γ in both cases such that total abatement of the grand coalition corresponds to our base case. This implies that for all values of total abatement smaller than the grand coalition abatement, the marginal benefit functions lie above the base case functions. Therefore, equilibrium abatement levels will be higher for any coalition smaller than the grand coalition, though changes are moderate. For instance, in the case,

$\alpha = 0.95$, we increase γ by a factor of 1.39, implying a total abatement level of 581 instead of 553 Mton per year for the singleton coalition structure (cf. Table 1).

Our findings are as follows. For $\alpha = 0.99$ we find no changes of the stable coalitions compared to our base case. For $\alpha = 0.95$ two coalitions that are stable in the base case disappear, but two new stable coalitions emerge. For instance, under equal sharing no stable non-trivial coalition remains (coalition {EET, CHN, IND} was stable in the base case). The largest stable coalition still consists of no more than four members and the emission sharing scheme remains to be the superior to all other sharing schemes in terms of welfare and abatement. Under the emission sharing scheme, coalition {USA, EET, EEX, CHN} achieves 36.9% of the welfare level in the grand coalition (base case, 35.8%).[12] The respective relative differences for other sharing schemes remain practically the same. Hence, we conclude that our results are fairly robust with respect to a change of functional form of the benefit functions. Of course, if we set α to still lower levels, we would observe more pronounced changes in the results. In this case, as pointed out above, changes are not only due to leakage, but due to much stronger level effects.

7. Conclusion

Greenhouse gas abatement is a global public good. It is hardly surprising that the implementation of the Kyoto protocol is hampered by adverse incentives of potential coalition partners, although large coalitions could create large global gains from cooperation. Due to the public good character of abatement, the very success of coalitions undermines their viability. The more abatement a coalition achieves, the stronger are the incentives to free-ride. This paper explores the role of surplus sharing for coalition stability. We identify stable coalitions for a set of surplus sharing rules. In particular, we examine equal sharing and sharing proportional to claims. The results show that some of the sharing schemes, for example when claims reflect historical responsibilities (inverse emissions), generate only small and ineffective coalitions ({EET, BRA} and {CHN, BRA}). These achieve only 0.5% and 2.4% of the potential surplus of globally optimal carbon abatement, respectively. In our set of rules, proportional sharing with emission claims performs best. The coalition {USA, EET, EEX, CHN} achieves about 36% of the potential surplus. Emission claims set the right incentives to get the large emitters with low marginal abatement costs 'into the boat'. As a general pattern, we observe that stable coalitions are formed by CHN and regions with the largest claims. This is because CHN provides low-cost abatement options and is therefore an attractive coalition partner for regions with large claims. Hence, CHN is joined by EEC under income claims, by EEX under population claims, by IND (and others) under ability-to-pay claims, by USA (and others) under emission claims, by BRA with inverse emission claims, and by USA under damage cost and abatement cost claims.

[12] The increase is due to the level effect – marginal benefits functions are above the base case functions.

This paper studies the performance of a set of given sharing rules that have been proposed in the debate on climate change policies. The task for subsequent research is to use these insights for the design of sharing rules which will stabilize larger and more successful coalitions. For the success of a coalition, it is important to get the regions with low abatement costs to join. But these will do little unless regions with high marginal damage costs are also joining. Only this would lead to a large scale internalization of the externalities from carbon emissions.

Finally, our results indicate that concerns for equity, taking ability to pay or historical responsibilities into account, may well be counterproductive and may lead to small and ineffective coalitions.

Acknowledgements

We wish to thank two anonymous referees for their detailed and inspiring comments and all members of the STACO project for support.

References

Altamirano-Cabrera, J.-C. and Finus, M. (2006) Permit trading and stability of climate agreements, *Journal of Applied Economics*, forthcoming.

Babiker, M.H., Reilly, J.M., Mayer, M., Eckaus, R.S., Wing I.S., and Hyman, R.C. (2001) The MIT Emissions Prediction and Policy Analysis (EPPA) Model: Revisions, Sensitivities and Comparisons Results, MIT Report No. 41, Joint Program on the Science and Policy of Global Change, MIT, Cambridge, MA.

Barrett, S. (1992) 'Acceptable' allocations of tradeable carbon emission entitlements in a global warming treaty, in United Nations (ed.) *Combating Global Warming: Study on a Global System of Tradable Carbon Emission Entitlements*, United Nations, New York, 85–113.

Barrett, S. (1994) Self-enforcing international environmental agreements, *Oxford Economic Papers*, **46**, 878–94.

Bloch, F. (2003) Non-cooperative models of coalition formation in games with spillovers, in C. Carraro (ed.) *The Endogenous Formation of Economic Coalitions*, Edward Elgar, Cheltenham, 35–79.

Bosello, F., Buchner, B., and Carraro, C. (2003) Equity, development and climate change control, *Journal of the European Economic Association*, **1**, 600–11.

Botteon, M. and Carraro, C. (1997) Burden-sharing and coalition stability in environmental negotiations with asymmetric countries, in C. Carraro (ed.) *International Environmental Negotiations: Strategic Policy Issues*, Edward Elgar, Cheltenham, 26–55.

Carraro, C. (2000) Roads towards international environmental agreements, in H. Siebert (ed.) *The Economics of International Environmental Problems*, Mohr Siebeck, Tübingen, 169–202.

Carraro, C. and Siniscalco, D. (1993) Strategies for the international protection of the environment, *Journal of Public Economics*, **52**, 309–28.

Chander, P. and Tulkens, H. (1995) A core-theoretic solution for the design of cooperative agreements on transfrontier pollution, *International Tax and Public Finance*, **2**, 279–93.

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

Ellerman, A.D. and Decaux, A. (1998) Analysis of post-Kyoto CO2 emissions trading using marginal abatement curves, MIT Report No. 40, Joint Program on the Science and Policy of Global Change, MIT, Cambridge, MA.

Eyckmans, J. and Finus, M. (2006) Coalition formation in a global warming game: how the design of protocols affects the success of environmental treaty-making, *Natural Resource Modeling*, forthcoming.

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

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

Finus, M. (2003) Stability and design of international environmental agreements: the case of transboundary pollution, in H. Folmer and T. Tietenberg (eds) *The International Yearbook of Environmental and Resource Economics 2003/2004*, Edward Elgar, Cheltenham, 82–158.

Finus, M., and Rundshagen, B. (2003) Endogenous coalition formation in global pollution control: a partition function approach, in C. Carraro (ed.), *Endogenous Formation of Economic Coalitions*, Edward Elgar, Cheltenham, 199–243.

Finus, M., van Ierland, E., and Dellink, R. (2006) Stability of climate coalitions in a cartel formation game, *Economics of Governance*, forthcoming.

Folmer, H., and von Mouche, P. (2000) Transboundary pollution and international cooperation, in T. Tietenberg and H. Folmer (eds), *The International Yearbook of Environmental and Resource Economics 2000/2001*, Edward Elgar, Cheltenham, 231–66.

Gosseries, A.P. (2003) Historical emissions and free-riding, in L. Meyer (ed.) *Justice in Time: Responding to Historical Injustice*, Nomos, Baden-Baden, 355–82.

Hoel, M. (1992) International environmental conventions: the case of uniform reductions of emissions, *Environmental and Resource Economics*, **2**, 141–59.

Kverndokk, S. (1995) Tradable CO_2 emission permits: initial distribution as a justice problem, *Environmental Values*, **4**, 129–48.

Mäler, K.-G. (1989) The acid rain game, in H. Folmer and E.C. van Ierland (eds), *Valuation Methods and Policy Making in Environmental Economics*, Elsevier, Amsterdam, 231–52.

Moulin, H. (1987) Equal or proportional division of a surplus, and other methods, *International Journal of Game Theory*, **16**, 161–86.

Na, S.-L. and Shin, H.S. (1998) International environmental agreements under uncertainty, *Oxford Economic Papers*, **50**, 173–85.

Nordhaus, W.D. (1997) *Managing the Global Commons: The Economics of Climate Change,* MIT Press, Cambridge, MA.

Pfingsten, A. (1991) Surplus sharing methods, *Mathematical Social Sciences*, **21**, 287–301.

Plambeck, E.L. and Hope, C. (1996) An updated valuation of the impacts of global warming, in N. Nakicenovic, W.D. Nordhaus, R. Richels, and F. Toth (eds) *Proceeding of the Workshop on Climate Change: Integrating Science, Economics and Policy,* IIASA, Laxenburg.

Rose, A. (1992) Equity considerations of tradeable carbon emission entitlements, in United Nations (ed.) *Combating Global Warming. Study on a Global System of Tradeable Carbon Emission Entitlements*, United Nations, New York, 55–83.

Rose, A., Stevens, B., Edmonds, J., and Wise, M. (1998) International equity and differentiation in global warming policy: an application to tradeable emissions permits, *Environmental and Resource Economics*, **12**, 25–51.

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

Tol, R.S.J. (2001) Climate coalitions in an integrated assessment model, *Computational Economics*, **18**, 159–72.

Weikard, H.-P. (2004) Who should receive the CO_2 emission permits?, in R. Döring and M. Rühs (eds) *Ökonomische Rationalität und praktische Vernunft – Gerechtigkeit, Ökologische Ökonomie und Naturschutz: Festschrift für Ulrich Hampicke,* Königshausen und Neumann, Würzburg, 71–82.

World Bank (2002) *World Development Indicators 2002*, CD-ROM edition, The World Bank, Washington, DC.