The Role of Emissions Trading and Burden Sharing in International Climate Agreements with Heterogeneous Countries

Michael Jakob[1], March 2012

Potsdam Institute for Climate Impact Research (PIK)

Abstract

This paper presents a model in which cooperation between heterogeneous countries can arise through pure self-interest. Emissions trading creates economic surplus by exploiting asymmetries, which can be distributed via an appropriate burden sharing scheme in a way that ensures that membership to the agreement is compatible with countries' incentives to join. While this mechanism improves upon the business-as-usual outcome, it does not solve the underlying collective action problem, such that social welfare falls short of the social optimum.

Keywords: International Environmental Agreements, Coalition Game, Emissions Trading, Burden Sharing

JEL classifications: C72, H41, Q54

[1] Corresponding author: jakob@pik-potsdam.de, Tel. ++49-331 288 2650, P.O.Box 601203, 14412 Potsdam, Germany, www.pik-potsdam.de

1. Introduction

Some prominent studies dealing with coalitions formation maintain that a global climate agreement can only sustain a low number of participants due to free-riding incentives (e.g. Carraro and Siniscalco 1993; Barrett 1994). These models are based on two fundamental assumptions: (i) all countries have identical costs of providing a global public good and derive identical benefits from it, and (ii) any coalition maximizes the benefits of its members. Yet, recent research which relaxes these assumptions draws more optimistic conclusions regarding the likelihood of a global climate treaty. First, with heterogeneous countries, an appropriate transfer scheme can significantly increase coalition size and bring the level of climate protection close to the global optimum by combining countries with a high willingness to pay for abatement of emissions and those with low mitigation costs (Carbone et al. 2009; Weikard et al. 2006; Carraro et al. 2006; Altamirano-Cabrera et al. 2008). Second, if the coalition adopts a less ambitious target than the one that would maximize its members' joint welfare, free-rider incentives as well as the costs of membership are reduced, and larger coalitions which achieve more stringent levels of abatement are feasible (Finus and Maus 2008).

Combining the above considerations, we present a simple analytical model in which heterogeneous countries interlinked by an emissions trading scheme decide on joining the coalition to provide a global public good contingent on a pre-determined contribution that only depends on their type. This set-up is particularly relevant in the case of a global climate agreement, in which industrialized countries (which can be expected to display a higher willingness to pay for climate change mitigation), provide transfer to finance mitigation in developing countries. We show how in such constellations, asymmetries among countries provide incentives for cooperation. Hence, our paper adopts a perspective similar to Barrett (2001), who demonstrates how side-payments can increase participation in a cooperative agreement by exploiting asymmetries between countries[2].

[2] Another recent study examining international environmental agreements among asymmetric countries is McGinty (2007), who considers a benefit-cost rule for the ex-post distribution of economic surplus after the abatement game. We differ from this contribution by (i) considering ex-ante allocations of emission permits in a trading scheme, (ii) allowing for any distribution of permits instead of restricting our analysis to one

By contrast, our main focus lies on the central role of the burden sharing scheme which is used to distribute emission permits among countries, and we highlight possible trade-offs between equity and efficiency regarding permit allocation.

This paper proceeds as follows: Section 2 presents the model and investigates (i) the complementarity between countries with high benefits and such with low abatement costs, and (ii) the nature of the equilibrium, It further derives an expression for the resulting coalition's welfare compared to (iii) the non-cooperative as well as (iv) the first best outcome as well as (v) the one that would obtain with a coalition maximizing the joint welfare of its members. Finally, (vi) it elaborates the role of burden sharing and the burden sharing scheme and demonstrated how equity considerations regarding the initial distribution of emission permits have the potential to impede the formation of a coalition. Section 3 discusses the policy implications of our results and concludes.

2. A coalition model of emissions trading among asymmetric countries

This section first presents the actors' payoff functions and identifies the business-as-usual as well as the socially optimal levels of abatement. It then extends the base model by introducing emissions trading combined with an associated burden sharing scheme to allocate responsibilities to abate emissions among countries. Finally, it discusses implications with regard to the question: is there a possible trade-off between an equitable allocation of emission permits and achieving the highest possible level of total welfare?

particular allocation scheme, and (iii) highlighting the potential trade-offs between equity and efficiency in the initial allocation of emission permits.

Costs and benefits

Let there be two types of countries – N and S – with linear benefits and quadratic abatement cost functions for the global public good ‘climate change mitigation’, labeled *e*. Each country bears the costs of its own provision of e_i, but benefits from mitigation provided by all countries:

$$b^{N,S}(e) = b^{N,S} \cdot \sum_i e_i \tag{1}$$

$$c^{N,S}(e) = \frac{1}{2} c^{N,S} e_{N,S}^2 \tag{2}$$

Net benefits (i.e. ‘welfare’) for each country are then simply given by the difference between benefits and costs:

$$W^{N,S} = b^{N,S} - c^{N,S}. \tag{3}$$

Further, let there be N_N and N_S countries of each type, respectively.

The business-as-usual outcome

Acting in isolation, i.e. without a mechanism to establish cooperation between countries (let’s call this the ‘business-as-usual case’), each country maximizes its individual net benefit by choosing its e_i such that marginal costs equal marginal benefits:

$$b^{N,S} = c^{N,S} \cdot e_{N,S}^{BAU} \Rightarrow e_{N,S}^{BAU} = b^{N,S} / c^{N,S} \tag{4}$$

Hence, total abatement in the business as usual case is given by:

$$e_{tot}^{BAU} = N_N \cdot e_N^{BAU} + N_S \cdot e_S^{BAU} = N_N \cdot b^N / c^N + N_S \cdot b^S / c^S \tag{5}$$

The social optimum

Summing up the net benefits of all countries yields the following expression for social welfare[3] :

$$W^{tot} = N_N \cdot \left[b^N \cdot (N_N e_N + N_S e_S) - \tfrac{1}{2} c^N e_N^2\right] + N_S \cdot \left[b^S \cdot (N_N e_N + N_S e_S) - \tfrac{1}{2} c^S e_S^2\right] \qquad (6)$$

From which we can easily derive the socially optimal abatement efforts for both regions:

$$e_N^{opt} = \frac{b^N N_N + b^S N_S}{c_N}, \text{ and}$$

$$e_S^{opt} = \frac{b^N N_N + b^S N_S}{c_S} \qquad (7)$$

These expressions are quite straightforward: they simply state that in the social optimum, the marginal costs of abating one additional unit of emissions (i.e. $c^{N,S} \cdot e_{N,S}^{opt}$) equal the associated marginal social benefits accruing to all countries (i.e. $b^N N_N + b^S N_S$).

Consequently, total abatement in the social optimum which maximizes total welfare is then given by:

$$e_{tot}^{opt} = N_N \cdot e_N^{opt} + N_S \cdot e_S^{opt} = (N_N / c_N + N_S / c_S) \cdot (b^N N_N + b^S N_S) \qquad (8)$$

[3] This expression implicitly assumes a utilitarian social welfare function, which is a standard assumption in the literature on coalition formation (cf. Barrett 1994)

Emissions trading

Let us now consider the case in which countries have the opportunity to enter into a global climate agreement, in which emissions trading (or a similar transfer scheme) is adopted, such that (i) marginal abatement costs across all countries which are members of the coalition[4] are equalized at permit price p and (ii) burden sharing between them is specified by a pre-determined contribution of o_N and o_S for each N-type and S-type country that is a member to the agreement[5]. The number of countries of each type participating in the agreement is denoted by n_N and n_S, respectively.

At permit price p, the abatement level of each country participating in the agreement is determined by the condition that the marginal costs of abatement ($c_{N,S} \cdot e^C_{N,S}$) equal the permit price. Hence, we can express the abatement level as a function of the permit price:

$$e^C_{N,S} = p / c_{N,S} . \tag{9}$$

From this expression, we can derive the carbon price which balances supply and demand, taking into account that the burden sharing schemes requires each N-type (S-type) country to abate o_N (o_S) units of emissions:

$$n_N \cdot p / c_N + n_S \cdot p / c_S = n_N \cdot o_N + n_S \cdot o_S , \tag{10}$$

which results in the following expression for the permit price p:

$$p = \frac{n_N o_N + n_S o_S}{n_N / c_N + n_S / c_S} \tag{11}$$

[4] To keep the analysis tractable, we restrict the discussion to the case of a single coalition. See e.g. Asheim et al. (2006) for a recent discussion of a model featuring several (regional) climate agreements

[5] Unlike Helm (2003), who examines a model in which self-interested countries endogenously choose their respective reduction commitments, we take these commitments as given by the respective allocation scheme and focus on how the interplay between gains from trade and burden sharing influences the decision to participate in the coalition

Adopting the notation $x = n_N / n_S$, and focusing on coalitions which include at least one country of each type[6] (i.e. $1/N_S \le x \le N_N$) the permit price can be rewritten as:

$$p = \frac{x \cdot (o_N c_N) \cdot c_S + (o_S c_S) \cdot c_N}{x \cdot c_S + c_N} \qquad (11')$$

In order to make our model relevant for the case of an international climate agreement, we adopt the following three assumptions.

<u>A1</u>: The benefits of N-type countries exceed those of S-type countries, i.e. $b_N > b_S$

<u>A2</u>: For both types of countries, abatement requirements under the burden sharing regime exceed abatement undertaken in the business-as-usual case, i.e. $o_{N,S} > b_{N,S} / c_{N,S}$

<u>A3</u>: N-type countries' marginal costs of meeting their reduction commitment by pure domestic mitigation are higher than that of S-type countries, i.e. $o_N c_N > o_S c_S$

These assumptions are quite straightforward. A1 simply ensures that there is heterogeneity between countries with regard to their benefits. A2 excludes those cases in which participation in the international climate agreement is trivially fulfilled, as the required reduction commitments do not go beyond the abatement that would be performed in isolation. Finally, according to A3, we only consider cases in which the marginal costs of meeting the reduction commitments specified by the burden sharing scheme by means of purely domestic abatement of emissions are higher for the country featuring the higher benefits. Due to differences in marginal abatement costs, opportunities to create economic surplus from emissions trading arise. As will be demonstrated below, A3 ensures that burden sharing is defined such that countries with a higher willingness to pay

[6] It is easy to show that only coalitions including a non-zero number of N- as well as S-type countries achieve abatement levels that exceed BAU abatement (combine Corollary 2 with the incentive compatibability conditions for p, namely (16) and (17))

for climate change mitigation (i.e. higher benefits) provide financial transfers by acquiring emission permits from countries featuring lower benefits.

Using these assumptions, we establish three corollaries that will be useful for the further analysis of the coalition game.

<u>Corollary 1</u>: $\frac{dp}{dx} > 0$*, i.e. the price of emission permits rises with the share of Northern countries and falls with the share of Southern countries in the coalition.*

Corollary 1 is verified by simply taking the derivative of (11'), keeping in mind A3. □

<u>Corollary 2</u>: $o_S c_S < p < o_N c_N$*, i.e. the price of emission permits has an upper as well as a lower bound, which are defined by the marginal abatement costs at the abatement level prescribed by the burden sharing scheme.*

Corollary 2 follows directly from calculation the limiting values of (11') for $x \to 0$ and $x \to \infty$ in combination with A3 and Corollary 1. □

<u>Corollary 3</u>: N-type countries' actual abatement is below the reduction commitments specified in the burden sharing scheme, such that emissions trading results in the provision of transfer payments to Southern countries. Conversely, S-type countries' actual abatement exceeds their reduction commitments, such that they receive revenues from emissions trading.

Corollary 3 follows directly from Assumption 3, which can be rewritten as $e_N^C - o_N = p / c_N - o_N < 0$ and $e_S^C - o_S = p / c_S - o_S > 0$. □

All three results are quite intuitive: according to A3, N-type countries display higher marginal abatement costs than S-type countries at the respective abatement level prescribed by the burden sharing scheme. Hence, a higher share of N-type countries necessarily must result in a higher equilibrium price of emission permits, as stated in Corollary 1. In addition, in a carbon market, the equilibrium price of emission permits has to settle somewhere between the highest and the lowest marginal abatement costs that would obtain if all abatement were performed domestically, as stated in Corollary 2. Finally, Corollary 3 states that, as S-type countries feature lower marginal abatement costs compared to N-type countries (at the respective abatement level prescribed by the burden sharing scheme), the former will be net sellers of emission permits.

Incentives to join the climate agreement

For the purpose of this paper, we regard an international climate agreement as a stable coalition of countries that meet their abatement requirement under the burden sharing rule (o_N, or o_S, respectively) by any combination of domestic abatement and emission trading. In the literature, stable coalitions are defined in terms of internal and external stability (cf. Carraro and Siniscalco 1993). Internal stability implies that no country that is a member of the coalition has an incentive to leave the coalition; external stability means that no country that is not a member of the coalition has an incentive to join. Formally, these incentives are describes by a function ϕ, which evaluates the net benefits of being part of the coalition against the net benefits enjoyed by non-members. If a country stays out of a coalition in which n_N and n_S countries (of each respective type) already participate, its welfare maximization problem results in abatement equal to its business-as-usual level[7]. It enjoys the benefits of abatement of the n_N and n_S countries which are part of the coalition, and the (N_N-n_N) and (N_S-n_S) abating at business-as-usual. Hence, non-members' welfare is given by benefits minus mitigation costs:

$$W_{non-member}^{N,S} = b^{N,S} \cdot \left[n_N o_N + n_S o_S + (N_N - n_N) \cdot \frac{b_N}{c_N} + (N_S - n_S) \cdot \frac{b_S}{c_S} \right] - \frac{1}{2} c^{N,S} (e_{N,S}^{BAU})^2 \quad (12)$$

[7] This is due to the linear benefit function which yields constant marginal benefits, such that country *i*'s marginal benefits from own abatement efforts are independent from all other countries' abatement

On the other hand, if the country joins the coalition, it (i) enjoys the additional benefits brought about by its own contribution to the coalitions, (ii) incurs domestic abatement costs for domestic abatement at price p, (i.e. $e^{C}_{N,S} = p / c_{N,S}$), and (iii) receives/provides transfer payments from emissions trading proportional to the difference between its reduction commitment and its domestic abatement (i.e. $p \cdot (e^{N,S} - o^{N,S})$).

Therefore, the benefits of joining a coalition with n_N and n_S members of N- and S-type, respectively, are:

$$W^{N}_{col} = b^{N} \cdot \left[(n_N + 1) \cdot o_N + n_S o_S + (N_N - n_N - 1) \cdot \frac{b_N}{c_N} + (N_S - n_S) \cdot \frac{b_S}{c_S} \right] - \tfrac{1}{2} \cdot c_N \cdot (\frac{p}{c_N})^2 + p \cdot (\frac{p}{c_N} - o_N) \tag{13}$$

$$W^{S}_{col} = b^{S} \cdot \left[n_N o_N + (n_S + 1) \cdot o_S + (N_N - n_N) \cdot \frac{b_N}{c_N} + (N_S - n_S - 1) \cdot \frac{b_S}{c_S} \right] - \tfrac{1}{2} \cdot c_S \cdot (\frac{p}{c_S})^2 + p \cdot (\frac{p}{c_S} - o_S) \tag{14}$$

The function ϕ, which describes the incentives to join the coalition compared to free-riding, is then given as the difference between $W^{N,S}_{col}$ and $W^{N,S}_{non-member}$:

$$\phi^{N,S} = b^{N,S} \cdot (o_{N,S} - b_{N,S} / c_{N,S}) - \tfrac{1}{2} \cdot c_{N,S} \cdot (p / c_{N,S})^2 + p \cdot (p / c_{N,S} - o_{N,S}) + \tfrac{1}{2} \cdot c_{N,S} \cdot (b_{N,S} / c_{N,S})^2 \tag{15}$$

The following corollary captures how countries' incentives to become a member of the coalition depend on the price of emission permits.

Corollary 4: Northern countries' incentives to join the coalition decline with rising prices, while for Southern countries they increase (i.e. $\frac{d\phi^N}{dp} < 0$ and $\frac{d\phi^S}{dp} > 0$).

Corollary 4 is rather intuitive, given that Northern countries are net importers and Southern ones net exporters of emission permits, as established in Corollary 3. Formally, it can easily be shown: $\frac{d\phi^{N,S}}{dp} = p / c_{N,S} - o_{N,S}$, which (in combination with A3) yields $\frac{d\phi^N}{dp} < 0$ and $\frac{d\phi^S}{dp} > 0$.□

We can now use the above observations to examine incentives for coalition membership, which allows us to determine the size and stability of the coalition. In particular, the following proposition establishes how the incentives for one type of country to participate depend on the participation of countries of the opposite type.

Proposition 1: A higher share of Northern countries in the coalition (i.e. a higher x) raises the carbon price. For Northern countries, this lowers the incentive to join the coalition, while it is increased by a higher share of Southern countries (which lowers the carbon price). Vice versa, for Southern countries, a higher share of Northern countries in the coalition (which raises the carbon price) increases the incentive to join, while it is lowered by a higher share of Southern countries (as it lowers the carbon price).

Proof: Proposition 1 follows directly from combining Corollaries 1 and 4, which can be combined to yield $\frac{d\phi^N}{dx} < 0$ *and* $\frac{d\phi^S}{dx} > 0$.. □

The central insight provided by Proposition 1 is that there is complementarity between N-type and S-type countries with regard to coalition membership: the incentives for each type of country to join the coalition are negatively affected by a higher share of countries of the same type in the coalition, while they are positively by a higher share of countries of the opposite type.

Coalition size and stability

We are now in a position to assess which stable coalitions can form in the game under study by examination of the incentive compatibility function ϕ. To start, we simplify the expression for ϕ by rewriting the incentive compatibility conditions for both types of countries:

$$\phi^N = \tfrac{1}{2} \cdot p^2 / c_N - p \cdot o_N + b_N \cdot o_N - \tfrac{1}{2} \cdot b_N{}^2 / c_N \tag{15'}$$

$$\phi^S = \tfrac{1}{2} \cdot p^2 / c_S - p \cdot o_S + b_S \cdot o_S - \tfrac{1}{2} \cdot b_S{}^2 / c_S \tag{15''}$$

Internal stability requires $\phi^N \geq 0$ and $\phi^S \geq 0$, i.e. that no country that is a member of the coalition has an incentive to leave. We now examine the range of permit prices for which N-type and S-types, respectively, prefer to be a member of the coalition instead of engaging in free-riding behavior.

First, solving (15') for feasible prices for which N-types countries have an incentive to join the coalition (keeping in mind A2) yields:

$$p_{1,2} \leq (c_N \cdot o_N) \pm (c_N \cdot o_N - b_N) \tag{16}$$

Due to Corollary 2, only the negative sign is admissible for the second term, such that $p \leq b_N$, i.e. North will not pay a price for emissions reductions that exceeds its marginal benefits of climate change mitigation.

Second, (15'') yields the range of permit prices for which S-type countries have an incentive to join the coalition:

$$p_{3,4} \geq (c_S \cdot o_S) \pm (c_S \cdot o_S - b_S) \tag{17}$$

Due to Corollary 2, only the positive part of the second term is admissible, such that $p \geq 2 \cdot (c_S \cdot o_S) - b_S$. In conjunction with A2, it also follows that $p \geq b_S$.

A coalition containing N-type as well as S-type countries is hence only internally stable if, on the one hand, the permit price is low enough to make it worthwhile for N-type countries to participate (16), but on the other hand low enough to make participation worthwhile for S-type countries (17). Combining (16) and (17) directly results in the range of permit prices for which the coalition exhibits internal stability:

$$b_N \geq p \geq 2 \cdot (c_S \cdot o_S) - b_S \tag{18}$$

Before proceeding, we establish a final corollary, which captures the relationship between the permit price, abatement levels, and welfare.

<u>Corollary 5</u>: *For all internally stable coalitions, the welfare of the agreement's members increase in p (and hence in the level of total abatement). The same holds for total welfare.*

According to the condition that marginal costs equal marginal benefit, the price that maximizes the coalition members' welfare is given by $p^C = n_N \cdot b_N + n_S \cdot b_S$. Likewise, the price that maximizes total welfare is $p^T = N_N \cdot b_N + N_S \cdot b_S$ [8]. For levels below p^C higher prices of emission permits (which according to (9) result in higher levels of total abatement) increase the coalition members' welfare as well as total welfare, i.e.

[8] Note that the expressions for p^C and p^T simply express the Samuelson (1954) rule. It is straightforward to derive the expressions for p^C and p^T as well as the derivative of welfare with respect to p from equations (13), (14), and (12).

$\frac{dW^C}{dp} > 0$ and $\frac{dW^{tot}}{dp} > 0$. Given (16), only coalitions with prices $p \le b_N$ are feasible in an internally stable coalition, such that the resulting price of emission permits will neither exceed p^C nor p^T. □

Inserting the incentive compatibility conditions (16) and (17) into the expression for p (11') enables us to derive the following participation constraints:

$$x \le x_{\max} = \frac{c_N(b_N - o_S c_S)}{c_S(o_N c_N - b_N)} \text{ (>0, by A2 and A3) and} \tag{19}$$

$$x \ge x_{\min} = \frac{c_N(o_S c_S - b_S)}{c_S(o_N c_N + b_S - 2 o_S c_S)} \text{ (>0, by A2 and A3)} \tag{20}$$

Internally stable coalitions exist when (19) and (20) are simultaneously satisfied, i.e. there are values of x ($1/N_S \le x \le N_N$) that meet both conditions. This is the case if:

$$\frac{c_N(o_S c_S - b_S)}{c_S(o_N c_N + b_S - 2 o_S c_S)} \le \frac{c_N(b_N - o_S c_S)}{c_S(o_N c_N - b_N)}, \tag{21}$$

which is equivalent to

$$b_N + b_S \ge 2 \cdot o_S \cdot c_S\,. \tag{22}$$

The participation constraints (19) and (20) now allow us to determine the size of internally stable coalitions in the following proposition.

<u>Proposition 2</u>*: Except for one special (degenerate) case, all internally stable coalitions contain all countries of at least one type (i.e. either all N-type or all S-type countries). If* $x_{\min} \le N_N / N_S \le x_{\max}$*, all countries will be member of the coalition.*

Proof: see Appendix.

The above proposition highlights one of the central arguments of this paper: with asymmetric countries, wide-spread participation or even universal participation in a global climate agreement is feasible. However, as we assume a coalition whose members commit to pre-defined reduction commitments instead of maximizing joint welfare, universal participation is not a sufficient condition for the socially optimal provision of climate change mitigation. Hence, the following two propositions state how the coalition performs compared to the business-as-usual case, and the social optimum, respectively.

Proposition 3: The total additional abatement that can be achieved with a coalition in place is the larger compared to the business-as-usual case (i) the larger the number of S-type countries, (ii) the larger the difference between the benefits of N-type and S-type countries, and (iii) the lower the abatement costs of S-types.

Proof: Maximum abatement occurs if there is a coalition in which the carbon price is at the maximum level which respect the incentive compatibility condition (18), i.e. $p = b_N$. According to Corollary 5, if the burden sharing is chosen in a way such that the coalition is internally stable[9], $p = b_N$ is the price resulting in the highest achievable level of the coalition members' welfare, as well as total welfare. Overall abatement then is $e_{tot}^{C} = N_N \cdot (b_N / c_N) + N_S \cdot (b_N / c_S)$, compared to $e_{tot}^{BAU} = N_N \cdot (b_N / c_N) + N_S \cdot (b_S / c_S)$ in the business-as-usual case. Hence, the maximum additional abatement achieved by cooperation then amounts to $e_{tot}^{C} - e_{tot}^{BAU} = N_S \cdot (b_N - b_S) / c_S$. □

Proposition 4: Abatement levels achievable with cooperation fall short of the social optimum. The difference is the larger (i) the larger the number of countries of each type, and (ii) the larger their respective benefits.

[9] As the proposition examines the maximal additional abatement that can be achieved, only internally stable coalitions need to be examined.

Proof: As we have demonstrated in Proposition 3, the maximum abatement that can be achieved is $e_{tot}^{C} = (N_N / c_N + N_S / c_S) \cdot b_N$, while the socially optimal level would be $e_{tot}^{OPT} = (N_N / c_N + N_S / c_S) \cdot (b_N N_N + b_S N_S)$. Hence, the maximum amount of climate change mitigation that is achievable with cooperation falls short of the social optimum by $e_{tot}^{OPT} - e_{tot}^{C} = (N_N / c_N + N_S / c_S) \cdot \left(b_N (N_N - 1) + b_S N_S\right)$. □

As pointed out in Proposition 4, emissions trading (in conjunction with a burden sharing scheme) does not solve the fundamental collective action problem. For this reason, the following proposition examines coalition size and stability if the coalition aims at maximizing the joint welfare of its members instead of basing reduction commitments on a pre-defined burden sharing scheme.

Proposition 5: If the coalition adopts the rule of joint welfare maximization instead of relying on exogenously given reduction commitments, no internally stable coalition can form.

Proof: Solving the coalition's welfare maximization problem yields the familiar condition that the emission price equals the coalition's marginal benefit: $p = n_N b_N + n_S b_S$. However, by (18), N-type countries only have an incentive to be in the coalition as long as $p \le b_N$. The price that would maximize the welfare of the coalition members violates the incentive compatibility condition for N-type countries. Hence, keeping in mind the restriction $n_N > 0$ and $n_S > 0$, no internally stable coalition is feasible. □

Finally, we analyze how countries' incentives to become members of the coalition depend on the burden sharing scheme adopted in the following proposition.

Proposition 6: The coalition's stability crucially depends on the burden sharing scheme adopted, i.e. the reduction commitments o_N and o_S allocated to N-type and S-type countries.

Proof: First, the participation constraint (22) establishes an upper bound for the maximum reduction commitment for which S-type countries have an incentive to join the coalition: $o_S \leq (b_N + b_S)/2c_S$. Second, the condition that $x > 1/N_S$ results in the upper bound for the reduction commitments of N-type countries: $o_N \leq b_N(1/c_N + N_S/c_S) - o_S$. Hence, internally stable coalitions that satisfy both participation constraints can be obtained by simply choosing appropriate reduction commitments o_N and o_S. More specifically, noting that for $o_S \to b_S/c_S$, $x_{\min} \to 0$ and for $o_N \to b_N/c_N$, $x_{\max} \to \infty$, a coalition featuring full membership can be obtained by appropriate allocation of reduction commitments. □

Besides demonstrating the importance of the burden sharing scheme adopted to distribute emission permits, Proposition 6 also bears important implications for climate policy. Universal participation in a global climate agreement can be achieved through an adequate initial distribution of emission permits. However, nothing guarantees that such a distribution, does not contradict fundamental equity considerations, such as distributing emission permits on an equal per-capita basis, or based on historical responsibility.

3. Discussion and Conclusions

In the literature on coalition formation, it has repeatedly been emphasized how acting out of self-interest results in free-riding behavior, so that cooperation is unlikely to occur. The model presented in this paper shows how cooperation can arise through pure self-interest. Three main conclusions can be derived. First, exploiting asymmetries between countries has the potential to improve upon the business-as-usual case, in which each country chooses its abatement level in isolation, even if countries act in a purely self-interested manner. In particular, we have highlighted how emissions trading in conjunction with an appropriate burden sharing scheme (defined by the allocation of emission permits among coalition members) can simultaneously (i) create economic surplus by equating marginal abatement costs across countries, and (ii) distribute the surplus such that for all countries it is in their best interest to cooperate. Second, our analysis also shows that this mechanism does not solve the underlying collective action problem, and that the resulting outcomes will always fall short of the social optimum, even if full participation is achieved. Third, incentives to participate in a global climate agreement crucially depend on the respective burden sharing scheme. With asymmetric countries (with regard to the benefits of climate change mitigation), it is always possible to choose a burden sharing scheme that ensures full participation in the global climate agreement. However, feasible burden sharing schemes that guarantee that all countries will have an incentive to join the coalition might turn out be fundamentally at odds with equity considerations, such as distributing emission permits on a equal per-capita basis, or based on historical responsibility for the already existing stock of greenhouse gases in the atmosphere due to past emissions (cf. Markandya 2011 for an overview of the relevant equity dimensions, and Bodansky 2004 for a summary of proposed burden sharing principles).

References

Altamirano-Cabrera, J.-C., Finus, M. and Dellink, R. (2008): Do Abatement Quotas Lead To More Successful Climate Coalitions? Manchester School, vol. 76(1), pp. 104-129

Barrett, S. (1994): Self-Enforcing International Environmental Agreements, Oxford Economic Papers, New Series, Vol. 46, Special Issue on Environmental Economics, pp. 878-894

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

Bodansky, D. (2004): International Climate Efforts Beyond 2012: A Survey of Approaches, Prepared for the Pew Center on Global Climate Change

Carbone, J., Helm, C., and Rutherford, T.F. (2009): The case for international emission trade in the absence of cooperative climate policy, Journal of Environmental Economics and Management, vol. 58(3), pp. 266-280

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

Carraro C. and D. Siniscalco (1993): Strategies for the International Protection of the Environment, Journal of Public Economics, vol. 52(3), pp. 309-328

Finus, M. and Maus, S. (2008): Modesty May Pay! Journal of Public Economic Theory, vol. 10(5), pp. 801-826

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

Markandya, Anil, 2011: Equity and Distributional Implications of Climate Change, World Development, vol. 39(6), pp. 1051-1060

McGinty, M. (2007): International environmental agreements among asymmetric nations, Oxford Economic Papers, vol. 59(1), pp. 45-62

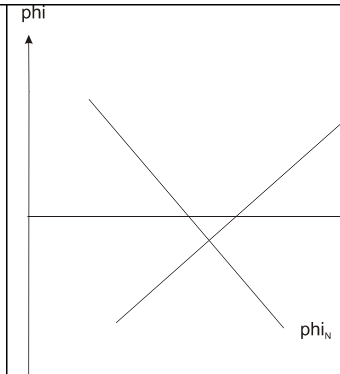
Weikard, H.P. Finus, M., and Altamirano-Cabrera, J.C. (2006): The impact of surplus sharing on the stability of international climate agreements, Oxford Economic Papers, vol. 58, pp. 209-232

Appendix

Proof of Proposition 2: From Corollaries 1 and 4, we can directly conclude that $\frac{d\phi_N}{dx} < 0$ and $\frac{d\phi_S}{dx} > 0$. The general behavior of the functions ϕ_N and ϕ_S is schematically depicted in Figure 1. Internally stable coalitions exist if neither N-type nor S-type countries have an incentive to leave the coalition, i.e. if there are values of x for which $\phi_N \geq 0 \vee \phi_S \geq 0$. Clearly, for $x_{\min} > x_{\max}$ as depicted in panel (a), the former condition is not fulfilled, and no internally stable coalition can emerge. On the other hand, for $x_{\min} < x_{\max}$ the emerging coalition is internally stable, but not externally (i.e. no x for which neither N nor S countries have incentives to join exist). Externally unstable coalitions can include all countries (panel b), all S-types, but not all N-types (panel c), or all N-types, but not all S-types (panel d). From any x for which $\phi^{N,S} > 0$, N- and/or S-type countries join the coalition until either no non-members are left, or until one of limits specified in (19) and (20) – i.e. $x = x_{\min}$ or $x = x_{\max}$ – is reached. Due to the complementarity between N and S, this can only occur in the case in which all countries of one type are coalition members. If $x_{\min} \leq N_N / N_S \leq x_{\max}$, the contingent of non-members is exhausted before one of the limits is reached (i.e. there would still be an incentive for countries of both types to join the coalition) and the coalition with universal participation emerges. Otherwise (i.e. if $N_N / N_S < x_{\min}$, or $N_N / N_S > x_{\max}$) countries will join the coalition until all countries of one type (N-type, or S-type, respectively) are members and countries of the other type have no more incentives to join (i.e. one of the limits is reached). Finally, for the special case $x_{\min} = x_{\max}$ (panel e), the coalition size is not uniquely determined; all coalitions for which $n_N / n_S = x_{\min}$ are (internally as well as externally) stable; these do not necessarily include all countries of either N-type or S-type. □

(a)
no coalition:

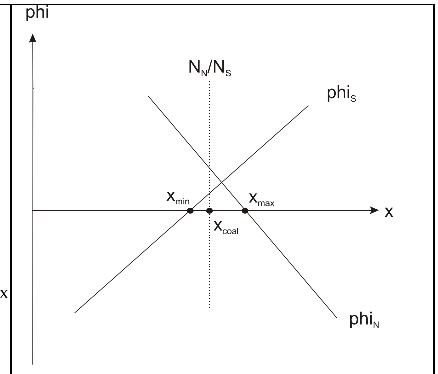
$x_{\min} > x_{\max}$



(b)
coalition with full participation:

$x_{\min} < x_{\max}$

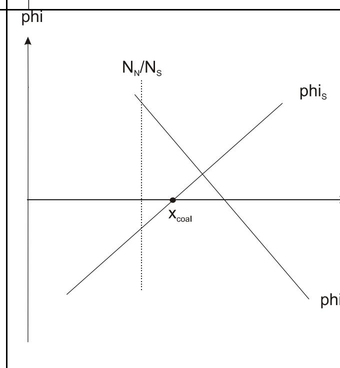
$x_{\min} \leq \frac{N_N}{N_S} \leq x_{\max}$



(c)
coalition including all S-type countries:

$x_{\min} < x_{\max}$

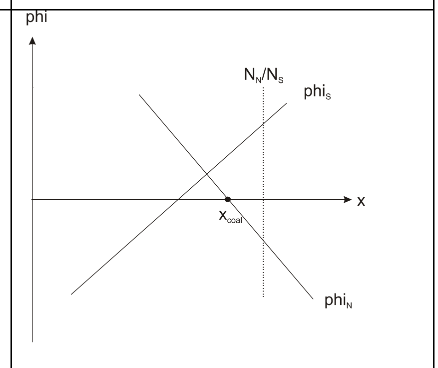
$\frac{N_N}{N_S} < x_{\min}$



(d)
coalition including all N-type countries:

$x_{\min} < x_{\max}$

$\frac{N_N}{N_S} > x_{\max}$



(e)
'degenerate' case:

$x_{\min} = x_{\max}$

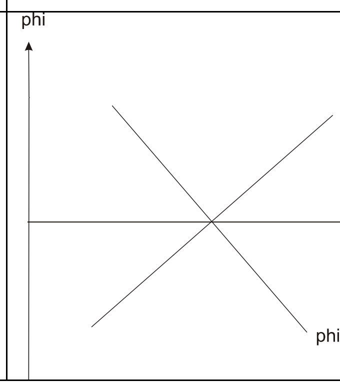


Figure 1: Possible outcomes of the coalition game