Citation for published version:
Lessmann, K, Marschinski, R, Finus, M, Kornek, U & Edenhofer, O 2014, 'Emissions trading with non-signatories in a climate agreement - an analysis of coalition stability', *The Manchester School*, vol. 82, no. S2, pp. 82–109. https://doi.org/10.1111/manc.12045

DOI:
10.1111/manc.12045

Publication date:
2014

Document Version
Early version, also known as pre-print

The Manchester School Vol •• No. •• ••–•• •• 2013
doi: 10.1111/manc.12045

EMISSIONS TRADING WITH NON-SIGNATORIES IN A CLIMATE AGREEMENT—AN ANALYSIS OF COALITION STABILITY*

by
KAI LESSMANN
Potsdam Institute for Climate Impact Research, Germany
ROBERT MARSCHINSKI
Potsdam Institute for Climate Impact Research, Germany
Technical University Berlin, Germany
MICHAEL FINUS
University of Bath
ULRIKE KORNEK
Potsdam Institute for Climate Impact Research, Germany
and
OTTMAR EDENHOFER†
Mercator Research Institute on Global Commons and Climate Change, Germany
Potsdam Institute for Climate Impact Research, Germany
Technical University Berlin, Germany

We investigate how different designs of carbon offset mechanisms, like the Kyoto Protocol's Clean Development Mechanism (CDM), affect the success of self-enforcing climate treaties. In a game-theoretic numerical model of coalition formation we find that participation in the agreement is negatively affected when strategic behavior and free-rider incentives matter. This does not change when selling targets restrict credit supply. Substantially higher participation emerges when the treaty restricts its signatories not to use the gains from credit trading to lower their emission caps. Despite the high sensitivity of participation to different CDM design, we find that global welfare levels achieved in various equilibria are remarkably similar.

1 Introduction

The Kyoto Protocol introduced three flexible mechanisms. The emission trading system (ETS) and joint implementation (JI) allow trading in emission entitlements among Annex I countries, those countries which accepted emission ceilings. In contrast, the Clean Development Mechanism (CDM) provides an opportunity for Annex I countries to buy emission credits from non-Annex I countries, i.e. mainly developing countries, which have not accepted emission ceilings. The CDM includes an additionality clause which requires that emission credits offered by non-Annex I countries must correspond to a reduction of emission levels 'below [what] would have occurred in

---

* Manuscript received 18.10.12; final version received 6.9.13.
† This work was in part funded by the German Federal Ministry for Education and Research (BMBF promotion references 01LA1121A), which is gratefully acknowledged.

the absence of the registered CDM-project activity' (UNFCCC, 2002, p. 43). All three flexible mechanisms provide opportunities to save abatement costs. This is in particular true for the CDM because the difference in marginal abatement costs between Annex I and non-Annex I countries is likely to be large. On the one hand, Annex I countries have to resort to increasingly costly abatement options to meet their emission caps. On the other hand, non-Annex I countries do not face such constraints on their emissions. Additionally, they typically face less steep marginal abatement cost functions compared with Annex I countries.[1] As compliance costs are a major obstacle for signing ambitious climate treaties, one is inclined to expect that all flexible mechanisms, and in particular the CDM, should have a positive effect on the incentive to sign a climate treaty. However, the question arises whether this conclusion is also true when departing from the assumption of a first-best world, explicitly considering strategic effects and the need for self-enforcing treaties due to the lack of a supranational enforcement power. Given the current efforts to negotiate a post-Kyoto agreement, it is therefore of great importance to understand how the design of offset mechanisms will affect participation and the success of future climate treaties.

In a strategic context, there are at least two reasons why a credit trading scheme between members and non-members of a climate treaty may not have the intended positive effect on the success of a climate treaty. First, the option of emission credit trading will affect equilibrium emissions of members and non-members. If abatement cost savings translate into more ambitious abatement targets of members and this is matched by less ambitious abatement targets by non-members, free-riding might become more attractive. This is an equilibrium effect associated with carbon leakage. Second, CDM trading implies efficiency improvements and therefore gains from trade, but these gains are in general unevenly distributed between members and non-members. When the CDM seller rather than the CDM buyer realizes most of the gains, the effect on the treaty may be negligible.

The first issue can be addressed by restricting the members' choices of emission allowances such that the gains from trade are not used for the implementation of more ambitious abatement targets. The second issue could be fixed through the implementation of selling targets.[2] If non-Annex I countries can only sell emission credits that correspond to emission reductions below baseline emissions, a share of the gains from trade can be shifted to the members, making it more attractive for members to stay in a climate treaty and/or for non-members to join a treaty. We investigate the different options in a systematic way.

---

[1] This is for instance illustrated by the marginal abatement cost curves from two integrated assessment models reported in Criqui *et al.* (1999).

[2] Selling targets (Kim and Baumert, 2002), similar to non-binding targets (Philibert, 2000) and no-lose targets (Meckling and Chung, 2009), specify an emission path relative to a baseline with the understanding that only below this emission path emission reductions can be sold as credits.

Our paper draws on two strands of literature. The first strand analyzes the stability of self-enforcing international environmental agreements. This literature goes back to Hoel (1992), Barrett (1994) and Carraro and Siniscalco (1993). Since then various departures from the standard model have been analyzed which include for instance trade in commodities (Eichner and Pethig, 2013a) and trade policies (Barrett, 1997; Lessmann *et al.*, 2009; Eichner and Pethig, 2013b), linking environmental agreements to negotiations on other issues in general (Folmer *et al.*, 1993) and to cooperation in research and development specifically (Carraro and Siniscalco, 1997; Botteon and Carraro, 1998; Lessmann and Edenhofer, 2011), a minimum participation clause (Carraro *et al.*, 2009; Weikard *et al.*, 2009), multiple agreements (Finus and Rundshagen, 2003; Asheim *et al.*, 2006; Eyckmans and Finus, 2006) and modest emission reductions (Barrett, 2002; Finus and Maus, 2008). The two papers closest to ours are Altamirano-Cabrera and Finus (2006) and Hoel and Schneider (1997). However, the first paper analyzes emissions trading only among coalition members, and not among coalition members and outsiders as we do. The second paper considers the possibility that coalition members buy additional emission reductions from non-members, even though Hoel and Schneider (1997) do not use the term CDM. But, as argued in Finus (2003, pp. 116–118), this paper suffers from a couple of conceptual shortcomings which by construction lead to smaller coalitions through the CDM.

The second strand of literature analyzes the strategic incentive under a permit trading scheme with endogenous choice of emission allowances,[3] but stability of treaties is tested in a rather simplistic way. In a stylized model, Helm (2003) compares the Nash equilibrium among individual countries in a scenario without trading to a scenario in which permit trading is anticipated. He shows that the effect of permit trading on global emissions is ambiguous: countries with steep damage-cost functions may abate more but countries with flat damage-cost functions may choose larger emission allowances. His results are driven by countries' anticipation of how their allowance choices influence the equilibrium permit price and hence their revenues from trade. Furthermore, he shows that even if the scenario with permit trade would reduce emissions globally, this scenario may be vetoed by individual countries because it makes some of them worse off. And, conversely, a scenario implying higher global emissions may be endorsed by all countries due to its welfare enhancing effect. However, in Helm (2003), the decision whether to participate in an agreement abstracts from strategic membership decisions and is only based on the concept of profitability, similar to the analysis conducted in Carbone *et al.* (2009), who base their analysis on a calibrated

---

[3] Additional considerations when several individual carbon markets exist are discussed in Fankhauser and Hepburn (2010). They analyze how different market designs influence the success of linking domestic carbon markets across borders in the face of economic and political obstacles. In the present paper, we assume a common perfect carbon market to be in place and study its strategic implications on coalition formation. Departures from this assumption are briefly discussed in Section 5.

CGE model. Moreover, in Carbone *et al.* (2009) and Helm (2003) the design of an agreement is different from our game-theoretic model of coalition formation. In our paper, members of an agreement cooperate and internalize the externality among themselves (though not with respect to outsiders). In contrast, in those two papers, countries entering an agreement just benefit from the possibility of permit trading, but all countries decide non-cooperatively on their emission allowances before trade takes place.

In what follows, in Section 2, we first discuss the set-up and develop an intuitive understanding of the main driving forces. Then we informally introduce our numerical model in Section 3 and provide the details in the Appendix. Section 4 reports and discusses our results, and Section 5 concludes.

2 Model: Setting, Policy Options and Driving Forces

2.1 Coalition Formation Game

The aim of this study is to investigate the impact of various designs of the CDM on the success of self-enforcing international environmental agreements. We follow the mainstream of the literature and model an agreement as a two-stage cartel formation game with N players. The game is solved by backwards induction. In the first stage, players decide on membership, i.e. whether to sign an agreement and hence become a coalition member (which we sometimes also call signatory), or to remain a non-member (to which we refer sometimes also as non-signatory), acting as a singleton. In the second stage, players decide on their economic strategies. In our model, this means that countries decide on the number of emission allowances they issue to their industry, taking into consideration the possibility of exporting excess allowances or importing allowances if they are short, as this follows from the market equilibrium in the CDM market.

In the second stage, we solve for a Nash equilibrium between the coalition and the remaining players, often termed *Partial Agreement Nash equilibrium* (PANE, Chander and Tulkens, 1995) in the specific context of a single coalition, a special case of the *social coalitional equilibrium* (Ichiishi, 1981). This implies that the coalition *de facto* acts as a single player, coordinating strategies such as to internalize the externalities among their members. Non-members simply maximize their own welfare. To determine the first stage, we apply the concept of cartel stability following d'Aspremont and Gabszewicz (1986). In equilibrium, players have no incentive to revise their membership strategy, given the strategies of other players. That is, a coalition is internally stable if no member has an incentive to leave and externally stable if no non-member wants to join the coalition. Note that for symmetric players, following Hoel (1992) and Finus and Maus (2008) stability can be compactly summarized by a stability function:

$$\Phi = W_{i \in S}(n) - W_{i \notin S}(n-1) \tag{1}$$

with a coalition of n symmetric members being internally stable if the stability function Φ is non-negative at n and externally stable if it is negative at $n + 1$ and where S denotes the coalition and W_i individual welfare of player i.

The details of the underlying economic model are further explained in Section 3 and all details are provided in the Appendix.

2.2 CDM Policy Designs

In our model, all regions can decrease their emissions by lowering the emission intensity of their production. In addition, coalition members have a second option for mitigating climate change: they can buy permits on the international CDM market. This is implemented through the choice of emission allowances for all regions. A region's actual emissions may exceed its allowances if the shortfall is matched by imported emission permits. Likewise, regions may export emission permits, selling surplus emission allowances by choosing lower emissions. In the analysis of different designs of offset mechanisms, we consider the following scenarios:

NT: As a benchmark, we consider the no-trade (NT) scenario without permit trade. This allows to explore the incremental effects of allowing for CDM trade; all discussions of relative effects will be related to the NT scenario. Note that we also sometimes refer to the non-cooperative equilibrium, which is different. The NT scenario allows for the possibility of coalition formation whereas the non-cooperative equilibrium corresponds to the 'all singletons coalition structure', i.e. no coalition has formed.

CDM/ut: This scenario assumes *unrestricted trading* of CDM credits as an integral part of the climate treaty. That is, the coalition maximizes welfare by choosing emission allowances, taking full account of the possibility of unrestricted CDM trade.

CDM/rt: The *restricted trade* CDM scenario assumes that CDM trading is included in the climate treaty to reduce implementation costs but allowance choices remain at their NT benchmark values. This design requires the ability of coalition members to commit to constrain emission allowances, similar to the commitment required for minimum participation clauses, burden sharing rules, or 'modest' abatement targets. Essentially, coalition members only make use of CDM trading to lower their abatement costs, but do not use the gains from trade to choose more ambitious emission targets compared with the NT scenario.[4]

Selling targets: We generalize the Kyoto Protocol's concept of additionality by introducing *selling targets* for the CDM supplier. A selling target specifies reductions relative to the NT baseline scenario that need to be achieved before any emission credits can be sold in the CDM market. We refer to a selling target below the NT baseline as being *stringent*, and a selling target

---

[4] See, for example, Courtois and Haeringer (2012).

above the baseline is said to produce *hot air*. Our default requirement of additionality corresponds to the special case of taking the NT baseline as the selling target, i.e. a selling target $sel = 0.0$. In contrast, a selling target $sel = 0.1$ would require 10 per cent additional emission reduction below the NT baseline.

2.3 General Effects of CDM Trading on Stability of Agreements

In this section, we briefly discuss some general effects of CDM trading on the stability of coalitions. As CDM trade may raise both the pay-off of members as well as that of non-members overall conclusions depend on the relative size of the effects. In the following, we decompose the overall effect into different driving forces, which, in equilibrium, work simultaneously in most scenarios.

We distinguish three groups of players: (a) members, (b) non-members and (c) a representative CDM supplier. Members choose their emission allowances cooperatively and if their actual emissions are in excess of their allowances, they can buy credits from the CDM supplier. Members who leave the agreement become non-members. Non-members choose their emission allowances non-cooperatively but cannot trade credits. Also the CDM supplier chooses emission allowances non-cooperatively and if the actual emissions fall short of allowances, he can sell credits to members. The CDM supplier does not make a membership decision, although he will only engage in credit trading if this improves his welfare position compared with the NT scenario.

Taken together, only the CDM supplier and coalition members sell and buy permits, but non-members can be indirectly affected through changes of equilibrium emissions. Three types of effects can be distinguished:

i. *Cost-effectiveness Effect*. Trade between coalition members and the CDM supplier occurs whenever the marginal abatement costs of the CDM supplier are lower than those of coalition members. Then, the coalition can substitute costly domestic abatement by cheaper CDM permits thus improving cost-effectiveness of their abatement. If total abatement efforts remain constant, non-signatories are unaffected by trade. Overall, it becomes more attractive to stay in the coalition and/or to join the coalition.
ii. *Ambition Effect*. If signatories anticipate the options of permit trading, the possibility of buying credits *de facto* shifts their marginal abatement cost curve downward. Consequently, in equilibrium, the coalition will increase its abatement efforts, choosing lower emission allowances. This will have a positive effect on signatories but the effect on non-signatories will be even greater since they will get the extra abatement at zero cost. Hence, the impact of the ambition effect on the size of stable coalitions will be most likely negative.

iii. *Leakage Effect*. For downward-sloping reaction functions in abatement space (because abatement levels are strategic substitutes), additional abatement by signatories will be partially offset by increased emissions of non-signatories, which is typically called carbon leakage. This undermines the benefits of signatories while saving costs of non-signatories. The leakage effect will clearly lead to smaller stable coalitions.

As the first effect has a positive impact, the third one a negative and the second one also most likely a negative impact on the stability of coalitions, a quantitative analysis is necessary to derive more conclusive results. This will be conducted with a numerical model which is described subsequently.

3 NUMERICAL MODEL

3.1 Basic Setting

We use an extended version of the numerical model MICA (Modeling International Climate Agreements) in our analysis, which builds on the multiregion optimal growth model with international trade presented in Lessmann *et al.* (2009). The most important extension concerns the trade of emission permits. A detailed description of the model can be found in the Appendix.

MICA is an optimal growth model of the Ramsey type with N world regions. Each region allocates income to either consumption or investment at every point in time. Regions maximize welfare, which is the net present value of utility, either as a singleton if they do not belong to the coalition or jointly if they are a coalition member. We assume a standard utility function, i.e. utility is increasing in per capita consumption with diminishing marginal utility and is discounted at the pure rate of time preference. Income stems from the production of a single good, assuming a neoclassical production function with capital and labor as factor inputs. Economic growth is driven by exogenous population growth as well as exogenously improving labor productivity.

Greenhouse gas emissions are modeled as a byproduct of economic activities. Total global emissions drive greenhouse gas concentration, which in turn determines the temperature increase relative to pre-industrial levels. The damage function, adapted from Nordhaus and Boyer (2000), translates global warming into negative economic impacts. Impacts can be reduced at the cost of investing in a generic mitigation option, which lowers the emission intensity of economic production.

An alternative way of meeting emission targets is to buy emission allowances from other regions. In accordance with the Kyoto Protocol, we impose two restrictions on emission credit trading between coalition members and outsiders. (i) Under the Kyoto Protocol, countries that provide CDM credits must be signatories of the protocol but without abatement commitment (i.e. they are non-Annex I countries) and conversely non-Annex I countries cannot offer CDM credits. This is why we distinguish between 'regular' non-members and a representative CDM supplier who can offer CDM credits to coalition members but who will never join the coalition. We assume that the CDM supplier has little own motivation to reduce emissions.[5] (ii) Following the Kyoto Protocol's additionality clause, we make it a default requirement that CDM credits represent true emission reductions (as opposed to so-called 'hot air'). Additionality is defined in relation to the NT scenario. We assume a perfectly competitive market of emission credits. Trade in goods is the means to finance imports of allowances. Goods from different regions are perfect substitutes.

3.2 Calibration

In this analysis, we restrict our attention to symmetric players as it is common practice in many stylized models of coalition formation (e.g. Ulph, 2004; Barrett, 2006; Carraro *et al.*, 2009).[6] This renders the analysis much simpler and in particular more transparent. Nevertheless, we calibrate the model such that aggregate values (e.g. total global emissions, economic output as well as greenhouse gas concentration and temperature increase) correspond to those of other climate-economy growth models, e.g. RICE-2010 (Nordhaus, 2010), REMIND-R (Leimbach *et al.*, 2010) or WITCH (Bosetti *et al.*, 2006). The model is run over 250 years in 10 year periods, but reported results relate to the first 100 years. For instance, in the business-as-usual scenario, which corresponds to the non-cooperative equilibrium with no CDM trade, average economic growth over the next century is approximately 2.4 per cent (cf. 2.2 per cent in RICE-2010), and CO_2 emissions rise from close to 8GtC in 2005 to about 20GtC in 2105 (cf. 7.8GtC and 19.5GtC in RICE-2010), triggering a temperature rise by 2.0°C in 2105 with climate change damages amounting to 6.1 per cent of economic output (cf. 2.8°C and 3.3 per cent in RICE-2010). In contrast, under full cooperative behavior (i.e. all climate change damages are internalized), global CO_2 emissions in 2105 are 13.8GtC; the associated increase in global mean temperature is 1.5°C with damages amounting to 4.1 per cent of economic output in that year (cf. 2.0°C and 2.3 per cent in RICE-2010).[7]

---

[5] Technically, we implement this by limiting this region's perceived climate change damages to a fraction (1 per cent) of its actual damages (cf. Section *Internalization of Damages* in the Appendix).

[6] More precisely, symmetry applies to signatories and non-signatories, but not to the CDM supplier for whom we assume much lower perception of damages (1 per cent of the damages other players face). Hence, in the NT scenario, the CDM supplier's emission allowances are much higher than those of all other players and hence his marginal abatement costs are much lower. Another asymmetry with respect to the CDM supplier is considered in Section 4.7 in the context of a sensitivity analysis.

[7] Data from RICE-2010 has been taken from Nordhaus (2010) and its supporting material if possible, and from the available spreadsheet version of the model otherwise.

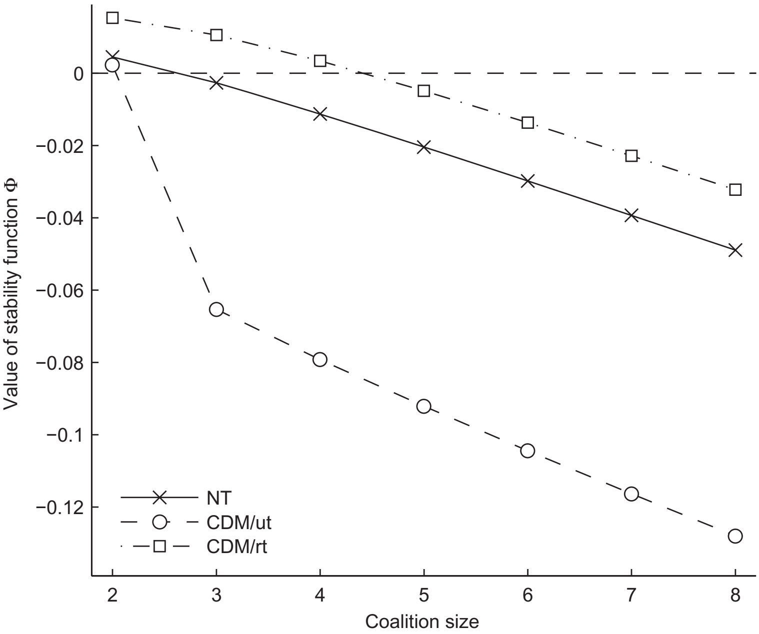

FIG. 1. Stability Functions with No CDM Trading (NT), with Unrestricted Trade of CDM Credits (CDM/ut) and with Restricted Credit Trade (CDM/rt)

4 RESULTS

4.1 NT Baseline

We begin with our benchmark, the NT scenario. The stability function Φ is shown in Fig. 1, which we may recall is the difference between the welfare of a player as a member in a coalition with n members, and the welfare when he leaves the coalition, becoming a non-member, and hence the coalition size is $n-1$ (equation (1)). As mentioned in Section 2.1, a coalition with n symmetric members is stable if the stability function is non-negative at n and negative at $n+1$. Thus, the stability function of the NT scenario indicates that only a coalition of two players is stable.

4.2 Unrestricted CDM Trading without Selling Targets (CDM/ut)

Figure 1 also shows how the coalition stability function changes when CDM trading is part of the agreement and unrestricted (denoted CDM/ut). Evidently, introducing CDM trading is counterproductive for participation as the stability function lies below the stability function of the NT baseline scenario, although in this specific example the actual size of the largest stable coalition still remains at 2, just as in the previous NT scenario.

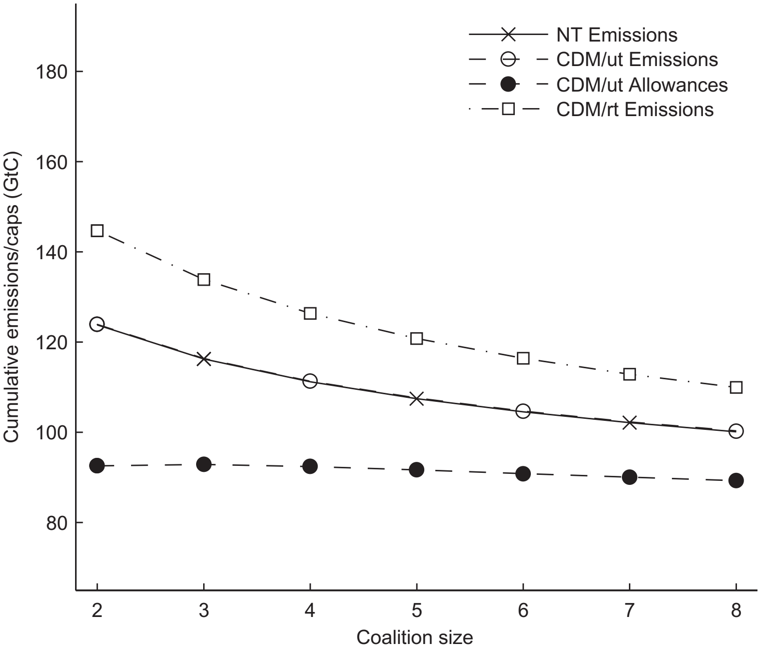

Fig. 2. Emissions of a Representative Coalition Member for Different Coalition Sizes (Cumulative Emissions over a Time Horizon of 100 Years)

The reason why stability decreases with CDM/ut is that the *ambition effect* identified in Section 2.3 outweighs the benefits from the *cost-effectiveness effect* accruing to the coalition (the third effect, the *leakage effect*, is less important). Figures 2 and 3 show the interaction of coalition members and CDM supplier in detail. Without CDM trade (NT scenario), the CDM supplier emits substantially more than coalition members. But in the CDM/ut scenario, the CDM supplier reduces emissions down to the same level as the one chosen by coalition members, in order to sell CDM credits: at this level, marginal abatement costs equal the permit price. The CDM credits are purchased by the coalition, but the lower abatement costs are mainly used to realize additional abatement rather than to replace costly domestic abatement by cheaper one from the CDM country (cf. member allowance choice in Fig. 2). This behavior of coalition members is driven by the shape of the marginal damage function, which is almost flat around the equilibrium. With (nearly) constant marginal damages, the coalition members simply keep on investing in abatement until marginal abatement costs reach the level of the sum of marginal damages of coalition members.

Overall, the abatement cost reduction from CDM trade results in lower global emissions. Thus, non-members benefit from lower damages. Although

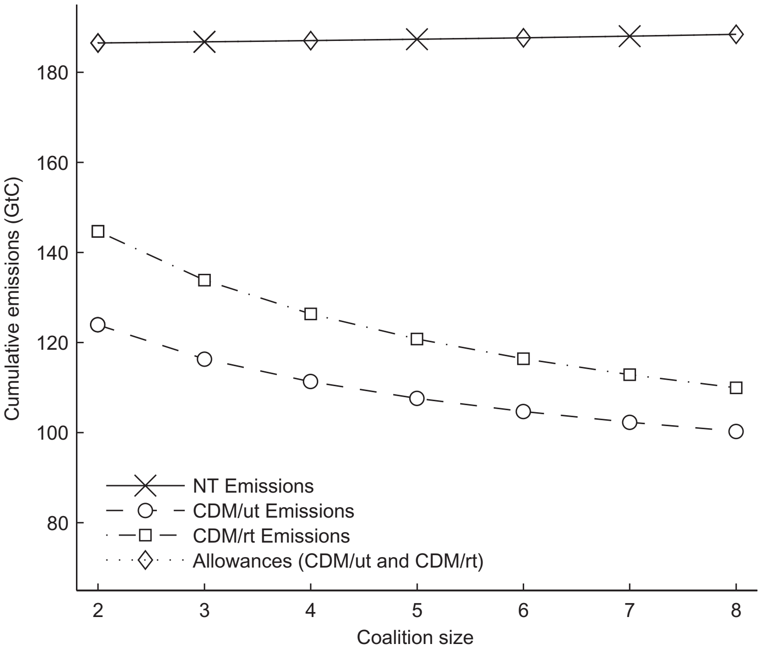

FIG. 3. Emissions and Allowances of the CDM Supplier for Different Coalition Sizes (Cumulative Emissions over a Time Horizon of 100 Years)

coalition members also benefit from lower damages due to the increased net abatement by the coalition, their abatement costs will generally rise due to the implementation of more ambitious abatement targets. Therefore, the gains from CDM trade are larger for non-members than for members and hence it becomes more attractive to leave a coalition of a given size as displayed in Fig. 1.

4.3 Unrestricted CDM Trade with Selling Targets (CDM/ut/sel)

Figure 4 provides an alternative illustration why CDM trade does not lead to larger coalitions. It shows the gains from CDM trade, i.e. the increase in global welfare as measured by the model's objective function, relative to welfare in the NT baseline scenario. These gains are unequally distributed among the different groups of players. For the default value of a zero selling target, which corresponds to the default additionality assumption, the welfare gains are appropriated by non-members and the CDM supplier. In particular, the CDM supplier benefits from selling emission allowances to coalition members. This is shown in Fig. 4 for a fixed coalition size of two members, but it also holds for other coalition sizes. This figure also shows that the coalition members suffer a slight loss of welfare relative to the NT

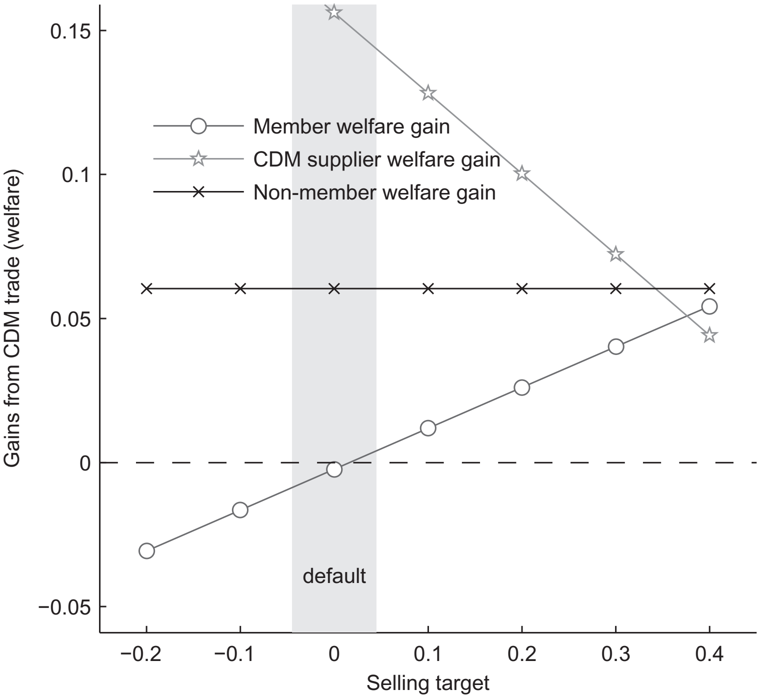

FIG. 4. Welfare Gains from CDM Trade (CDM/ut/sel) Relative to the NT Scenario for a Representative Coalition of Two Players

scenario. This is a side-effect of leakage: in equilibrium, members buy credits at their marginal damages, but due to leakage the actual reduction of global emissions is less. For small coalition, this effects dominates the efficiency gains from trade, for larger coalitions, this leakage effect vanishes.

Figure 4 also visualizes how the gains from CDM trade may be shifted from the CDM supplier to coalition members using selling targets. Selling targets specify emission reductions for the CDM supplier relative to his baseline emissions projected for this particular coalition size (in our case this is the NT scenario) that they need to achieve before they are allowed to sell further reductions into the CDM market. For instance, a selling target of 0.2 implies that the CDM supplier has to reduce 20 per cent compared with baseline emissions before selling emission credits.

More stringent selling targets shift welfare gains from the CDM supplier to coalition members. Essentially, by imposing selling targets, the coalition receives an emission reduction up to the selling target for free and only pays for additional emission reductions beyond the target. These gains come at the expense of the CDM supplier. In equilibrium, global levels of welfare and emissions remain constant for all selling targets because the selling targets do

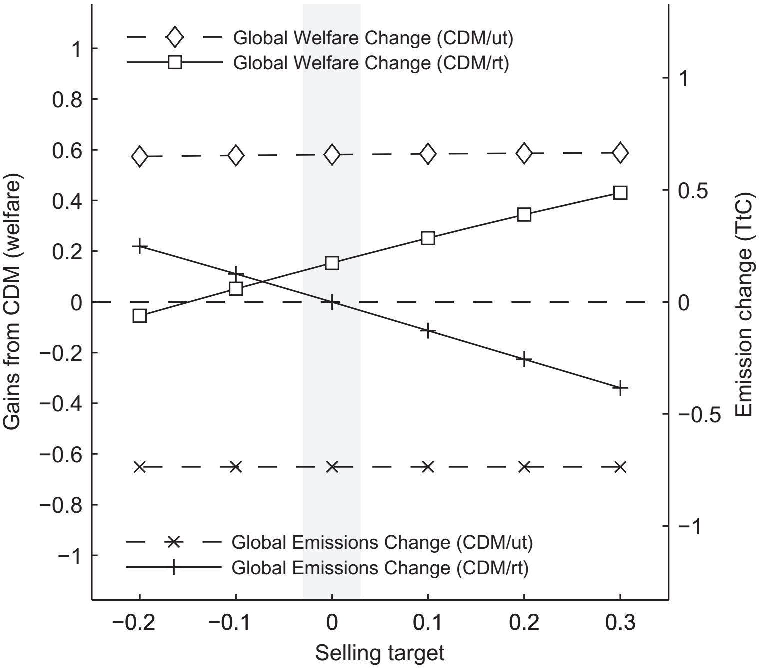

Fig. 5. Changes of Global Emissions and Global Welfare When Going from the NT Scenario to the CDM Scenarios for Different Selling Targets. Numbers are for a Representative Coalition of Four Players, and Welfare Gains are Expressed as the Difference in Aggregate Welfare between the Scenarios

not alter marginal abatement costs and marginal damages, which in turn determine the equilibrium allocation of abatement.[8] This is evident from Fig. 5, which shows the change in global total emissions (in teratons of carbon) and welfare (in terms of social welfare as defined by the model's objective function) brought about by CDM trade: positive numbers indicate that CDM trade raises emissions (or welfare).

In view of the fact that selling targets improve welfare of coalition members, a positive effect of selling targets on stability is very plausible. Indeed, the stability function for CDM/ut in Fig. 1 would shift upwards with more stringent selling targets (not shown). However, again, in this example, the shift is not sufficient to generate stable coalitions larger than two members. The largest stable coalition therefore remains unchanged.

In summary, the first type of offset design in the form of adding the CDM to the coalition agreement has a negative impact on coalition stability:

---

[8] Similar to the findings in Manne and Stephan (2005), a separability of efficiency and equity (i.e. the global least-cost allocation of abatement between countries and the distribution of the associated abatement costs) holds in MICA due to the feature of international trade in goods.

the benefits from CDM trade are realized on the side of the CDM supplier rather than on the side of coalition members; more importantly, non-members' welfare is increased, which raises the incentive to free-ride. Selling targets allow to counteract this effect. However, in our model, this is not sufficient to raise participation above the NT benchmark.

4.4 Restricted CDM Trade without Selling Targets (CDM/rt)

In the previous section it became clear that unrestricted CDM credit trading exacerbates the free-riding problem because coalition members have incentives to abate in excess of the NT scenario. The alternative design of restricted CDM trade tries to remedy this problem. CDM trade is introduced solely to reduce compliance costs for a given level of allowance choices. Essentially, this places a constraint on coalition members' emission allowances such that they cannot fall below their emissions in the NT scenario. Hence, we reduce the leakage effect and avoid that non-members benefit from additional abatement by the coalition.

Analogously to the CDM/ut scenario, the CDM/rt scenario is illustrated in Figs 1–3 and 5. In Fig. 1, we see our intuition about the overall effect of this CDM design confirmed: Coalition stability is improved such that a coalition of four players becomes stable. For a given coalition size, and for the default additionality assumption (i.e. a zero selling target) global welfare gains in the CDM/rt scenario fall short of those in the CDM/ut. This is because the additional constraint in CDM/rt prevents further abatement by the coalition compared with the NT scenario. Furthermore, we see that due to the additionality clause and the *restricted trade* setting, equilibrium allowances of members and CDM supplier correspond to their NT baseline emissions, and their emissions are higher than under CDM/ut (Figs 2 and 3) for a given coalition size. We now turn to explore whether selling targets could shift the stability function even further upward as they did in the CDM/ut scenario.

4.5 Restricted CDM Trade with Selling Targets (CDM/rt/sel)

The effect of CDM with selling targets on global emissions and welfare (for a given coalition) is shown in Fig. 5. Selling targets require additional abatement from the CDM supplier before he can sell credits in the market, and the CDM/rt assumption effectively establishes a lower bound for abatement by the coalition. Therefore, selling targets reduce global emissions. Since this moves global emissions closer to the social optimum, it has a positive impact on global welfare. This is quite contrary to CDM/ut where selling targets only redistributed welfare gains among players (i.e. between CDM supplier and coalition members), leaving the global levels of welfare and emissions untouched. Also in contrast to CDM/ut where members' gains increase and non-members' gains remain constant through an increase in selling targets

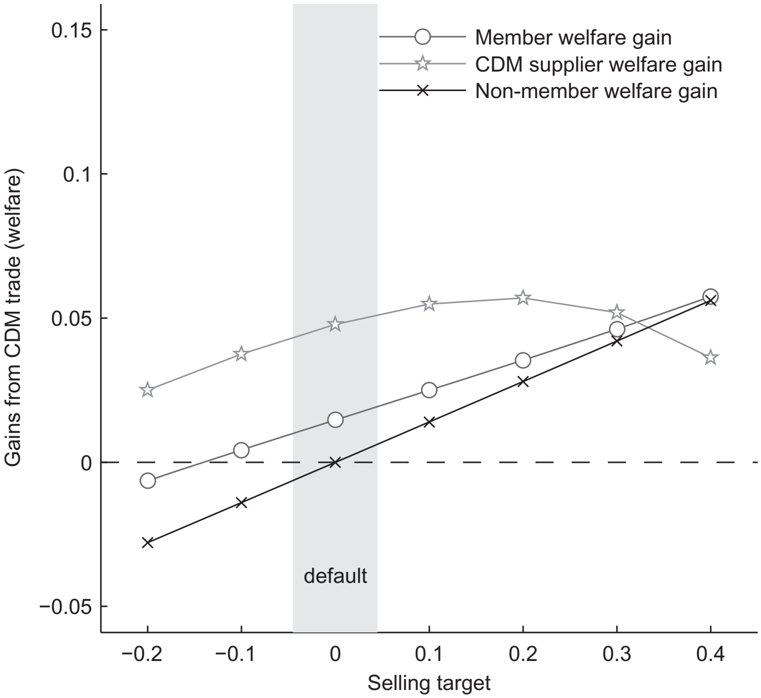

FIG. 6. Welfare Gains from CDM Trade (CDM/rt/sel) Relative to the NT Scenario Assuming a Representative Stable Coalition of Four Players

(Fig. 4), now members' and non-members' gains increase, but the difference between both declines with more stringent selling targets (Fig. 6), which may negatively affect the stability of coalitions.

The intuition for this is as follows: For CDM/rt, members and non-members benefit from globally reduced emissions. Members also benefit from a bounty of cheap CDM credits but with more stringent selling targets, credits become scarcer and more expensive, thus diminishing this benefit. For negative selling targets (i.e. hot air) this trend is reversed. While the world may be worse off with negative selling targets for a given coalition size (Fig. 5) and likewise, also members and non-members (Fig. 6), the negative effect on non-members exceeds that on members. Both are negatively affected by higher global emissions. However, for coalition members this is partially offset by the greater amount of CDM credits which are now available at a lower price, leading to a stabilization of a coalition of five members for a selling target of –0.2. Thus, there is a trade-off between environmental effectiveness and participation. The effect is similar to the idea of 'modest' emission reductions analyzed in Finus and Maus (2008). Given this trade-off, it is therefore important to have a look at the overall effects.

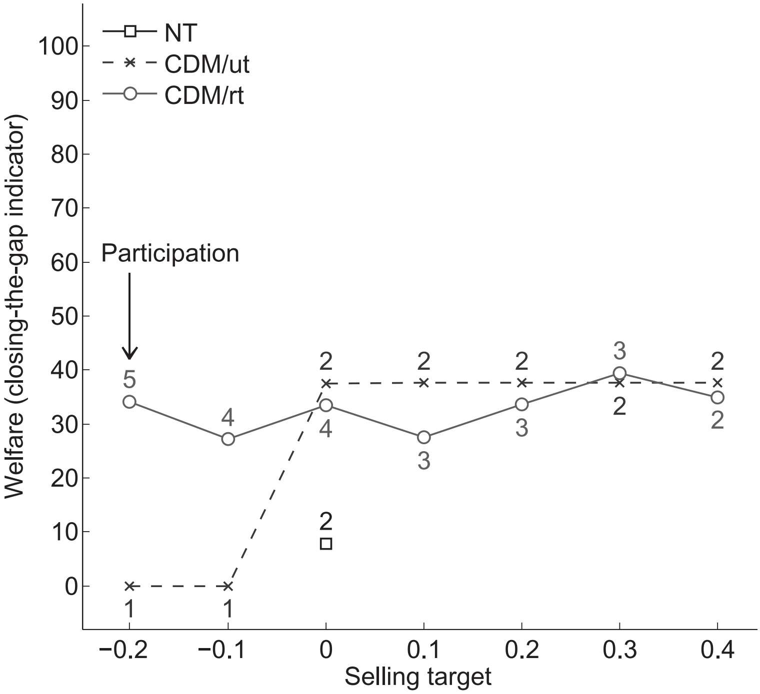

FIG. 7. Participation and Global Welfare for Different Selling Targets. Welfare is Scaled to the Gap between No Cooperation (0 per cent) and Full Cooperation (100 per cent)

4.6 *The Overall Effect of Selling Targets on Global Welfare*

Until now we analyzed the impact of various CDM trading scenarios on the size of stable coalitions. We also considered for a given coalition size how selling targets affect the welfare of the three groups in our model, as well as overall welfare. The impact on individual welfare was useful to understand how selling targets may change the incentive to leave or join the coalition. The impact on global welfare was useful in order to understand whether selling targets could be desirable from a normative point of view. The previous discussion revealed that effects are not always straightforward and that there may be trade-offs between the size of stable coalitions and global emissions. Hence, we now look at the overall effect.

Figure 7 summarizes the effect of credit trading on participation in the agreement and global welfare. Participation is indicated by the numbers next to the bullets. The level of welfare achieved in the presence of a given stable coalitions is scaled from 0 per cent (non-cooperative equilibrium) to 100 per cent (full cooperation, social optimum). Under the NT scenario, a coalition of two members achieves a welfare level of about 10 per cent.

Under CDM/rt, we find a positive effect of CDM credit trading on participation and global welfare compared with the NT scenario. For CDM/rt, welfare increases to levels of about 35–40 per cent and membership ranges between 2 and 5 members. Selling targets have not much impact on global welfare because more stringent selling targets reduce participation, although for a given coalition size they increase welfare as discussed above. In fact, as is evident from Fig. 7, the two opposing effects more or less offset each other.

The CDM/ut scenario is much simpler. The largest stable coalition always has two members, unless 'hot air' undermines stability completely. Through trading, global welfare increases compared with the NT scenario. However, as discussed previously, the global welfare achieved by a two-player coalition is independent of selling targets because selling targets only redistribute welfare.

Taken together, both CDM/ut and CDM/rt improve upon the NT baseline, with both achieving similar levels of global welfare, irrespective of selling targets.

4.7 Sensitivity Analysis

In the previous analysis, we only considered the case of symmetric players in order to focus on strategic effects of treaty design without having to deal with the complications that asymmetry entails. The only departure from symmetry relates to the CDM supplier for which we assumed a substantially lower anticipation of climate change damages. This implied that the CDM supplier's willingness to abate out of self-interest in the non-cooperative and the NT scenario is much lower than that of all other players. As a consequence, the CDM supplier has much lower marginal abatement costs before trade takes place. In this section, we explore a variation of this approach: we start with the CDM supplier being perfectly symmetric to all other players. Then, we gradually reduce the CDM suppliers marginal abatement costs (by shifting them downward) in order to increase the supply of CDM credits.[9] To keep everything else as much the same as possible (i.e. same overall equilibrium emissions compared with symmetry), and to be able to discuss the effects in a *ceteris paribus* manner, we simultaneously reduce the CDM supplier's marginal damages (by shifting them downward).[10] This means equilibrium emissions in the CDM supplier region are about the same as before as long as there is no CDM trade, but once trade is considered more CDM credits are on offer.[11]

Figure 8 summarizes the results. Four main conclusions are important. First, CDM trading improves upon the NT scenario. Second, restricted and

---

[9] In terms of the model equations given in the Appendix, this is implemented by increasing the investment effectiveness parameter ξ_i in equation (A8).

[10] Implemented by reducing parameter θ_{1i} of the damage function in equation (A14).

[11] We owe this scenario to an anonymous referee.

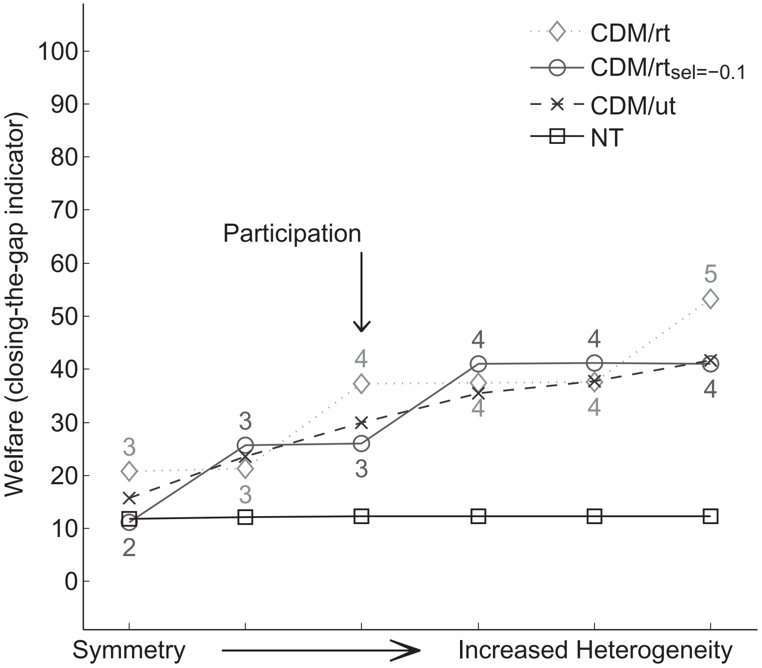

FIG. 8. Participation and Global Welfare for Different Degrees of Heterogeneity. Welfare is Scaled to the Gap between No Cooperation (0 per cent) and Full Cooperation (100 per cent). Participation for CDM/ut and NT is Always 2

unrestricted trade do not make much of a difference in terms of the overall outcome. Third, for restricted trade, there is a trade-off between participation and welfare for different selling targets. If selling targets do not affect membership, they improve global welfare. But selling targets may lead to lower participation. Overall, selling targets cannot really improve the overall welfare of restricted trade, and are anyway irrelevant for unrestricted trade as argued above. These three conclusions show the robustness of our conclusions obtained above. Fourth, due to the fact that we can now model the volume in CDM trade through a variation of heterogeneity between the CDM supplier and the remaining players, a new facet comes into play. The larger the degree of heterogeneity and hence the larger the volume in CDM trade, the larger will be participation and the larger the relative welfare gains in a stable agreement. Also the difference in relative performance of an agreement with CDM trade (CDM/ut and CDM/rt) compared with one without (NT) increases. This last conclusion is in line with intuition: if CDM trade can improve upon the success of treaty formation in a strategic context at all, then this improvement will be particularly large for large trade volumes.

5 CONCLUSION

This paper investigated how the success of a self-enforcing climate agreement is affected by emissions trading between Annex I and non-Annex I members of a climate agreement, as for instance the CDM of the Kyoto Protocol. In a first-best world, the CDM will clearly have an unequivocally positive effect, as it lowers total abatement cost. However, in a world with strategic interactions between countries and free-riding behavior, this is less evident. The interplay of three effects of the CDM—the *cost-effectiveness*, *ambition* and *leakage* effects—will determine the overall impact on international cooperation, requiring a detailed analysis based on numerical simulations. For instance, if the gains from CDM trade are higher for non-members than for members of a climate agreement, participation will actually be discouraged.

For the design of emission credit trading without any restriction (CDM/ut), we have shown that a negative impact on participation in the climate agreement should be expected. This can be understood as the unfavorable domination of the *ambition effect* over the *cost-effectiveness effect*: the availability of low-cost CDM abatement is used by the coalition to implement additional reductions and achieve lower global emissions, from which all countries benefit, but only the coalition members bear the costs. But while the *ambition effect* adversely affects the incentive to participate, any coalition that is stable despite this will achieve more emissions abatement than the comparable coalitions without CDM (NT) or with restricted trade (CDM/rt). Thus, CDM/ut tends to produce narrow-but-deep climate agreements.

Putting an additional constraint on the CDM supplier country by means of selling targets (CDM/ut/sel) allows to shift a larger share of the gains from trade of the CDM towards coalition members. Hence, we should expect that this improves membership in an agreement. However, in our model this effect turned out to be too weak to have a significant positive influence on the performance of the unrestricted CDM.

In view of this negative result, we investigated the alternative CDM design of restricted trade (CDM/rt) which excludes the possibility that the coalition countries use CDM trade to achieve more ambitious reduction targets. And even though for a given coalition size this implies higher global emissions and hence lower global welfare than unrestricted trade, our model also showed that this type of CDM has indeed a positive impact on participation, leading to larger stable coalitions compared with both the no CDM (NT) and unrestricted CDM (CDM/ut) cases. Compared with the stable coalitions of the unrestricted CDM design, stable agreements under CDM/rt are hence broad-but-shallow.

If, in addition, selling targets are introduced under this type of CDM (CDM/rt/sel), global welfare can be further increased since the CDM supplier now has to carry out additional abatement (at her own cost) before selling credits on the CDM market. However, it turns out that the stringency of the selling target has an ambivalent effect: although a more demanding selling target always implies higher global welfare for a *given* coalition size, the size of the largest coalition that is stable actually becomes smaller. As a consequence, a converse reasoning applies to the role of hot air (i.e. negative selling targets): while reducing the environmental effectiveness of an agreement, which clearly leads to reduced global welfare levels, it may at the same time help to draw additional members into the coalition. This is because it becomes less costly to comply with the watered down agreement. However, our simulations showed that the resulting larger coalitions were hardly able to outperform smaller coalitions without hot air.

Finally, the analysis of the aggregate impact of the different designs of the CDM on global welfare levels confirmed—as perhaps one of the most important result of our paper—that both the unrestricted (CDM/ut) and restricted (CDM/rt) type of CDM lead to significant welfare gains as compared with the no-trade base case (NT) (as long as the amount of hot air possibly introduced in CDM/rt remains small). When the resulting stable coalitions are compared across designs, it turns out that both CDM approaches lead to similar global welfare levels, despite the differences in participation levels. This remains true for all selling targets and any degree of heterogeneity we considered. Heterogeneity does, however, make a difference for the additional global welfare attainable through CDM: the larger the asymmetry between CDM seller and buyers, the higher the volume of CDM trade, and hence the higher the gains from agreement formation.

In summary, how an offset mechanism like the CDM is incorporated into an agreement has far reaching consequences concerning its participation, the agreement's ambition, and its distributional implications. Yet, the net effect in terms of the global welfare achieved when taking all strategic interaction into account is surprisingly similar for all the options considered here. Nevertheless, if high participation is seen as a value in itself, e.g. when the stable agreement is understood as the part of a process where climate policy is first 'broad, then deep' (Schmalensee, 1998), this could tip the scales towards the CDM/rt design. However, a full formal analysis to investigate this point comprehensively would require a model of dynamic membership and is beyond the scope of this paper.

While our model also shares many restrictions of stylized models, we think that one of the most interesting extension for future research is this consideration of dynamic membership. That is, whereas in our model the membership is a one-shot decision based on discounted utility, one could allow for the possibility that countries can revise their decision continuously as in Rubio and Ulph (2007). This would allow to study how the design of emission credit schemes affects participation in successive climate agreements. It might be that the possibility of offering CDM credits may not pay in the long run if current CDM suppliers do not accept emission ceilings in a future climate treaty simply because being a non-Annex I country and CDM supplier is more attractive than becoming an Annex I country. Another interesting aspect would be the impact of the co-existence of and links between different permit trading schemes, as currently observed in various regions, like the EU, the USA and Australia, on the formation of climate treaties. Such an analysis would require to consider coalition formation games with multiple coalitions as for instance analyzed in Finus (2008) and Finus and Rundshagen (2003), but such an extension would certainly be non-trivial in our rich policy setting.

APPENDIX: MODEL EQUATIONS

In this section, we present the details of our numerical model. The model builds on Lessmann *et al.* (2009) and Lessmann and Edenhofer (2011) and is extended to include the endogenous choice and trade of emission allowances. In the following, we first describe the model equations, their calibration, and the numerical procedure to solve the model.

TABLE A1
PARAMETERS AND INITIAL VALUES

Parameter	*Symbol*	*Value*
Rate of labor efficiency improvement	α	0.023
Income share capital	β	0.35
Abatement cost exponent	γ	0.2
Emission/concentration conversion factor	ζ	0.47
Elasticity of marginal utility	η	1
Damage function coefficient	θ_{1i}	0.02
Damage function exponent	θ_2	1.5
Rate of ocean CO_2 uptake	κ	2.15e-2
Labor efficiency	λ	$e^{\alpha t}$
Radiative temperature driving factor	μ	8.7e-2
Exogenous rate of decarbonization	ν	0.01
Effectiveness of investments in a_{it}	ξ_i	5.0
Pure rate of time preference	ρ	0.01
Temperature damping factor	ϕ	1.7e-2
Atmospheric retention factor	ψ	1.51e-3
Initial labor productivity	a_0	1
Initial concentration	C_0	377
Initial cumulative emissions	E_0	501
Initial capital stock	k_0	70
Initial labor	l_0	6.6
Initial temperature change	T_0	0.41

Preferences

The world economy is modeled as a set of $N = 9$ regions (or players). Players decide in an intertemporal setting which share of income to consume today and which share to save and invest for future consumption. Intertemporal welfare W_i and instantaneous utility function U, which is based on per capita consumption, are given by

$$W_i = \int_0^\infty l_{it} U(c_{it}/l_{it}) e^{-\rho t} \mathrm{d}t \tag{A1}$$

$$U(c_{it}/l_{it}) = \begin{cases} \dfrac{(c_{it}/l_{it})^{1-\eta}}{1-\eta} & \text{if } \eta \neq 1 \\ \log(c_{it}/l_{it}) & \text{if } \eta = 1 \end{cases} \tag{A2}$$

where c_{it} and l_{it} denote consumption and labor in region i at time t respectively. Parameter ρ is the pure rate of time preference, and parameter η denotes the elasticity of marginal utility.

Technology

The economic output y_{it} in each region is produced with a Cobb–Douglas production technology F with a capital income share of β. Climate change damages (to be defined below in equation (A14)) destroy a fraction $1 - \Omega_{it}$ of the production.

$$y_{it} = \Omega_{it} F(k_{it}, l_{it}) \tag{A3}$$

$$F(l_{it}, k_{it}) = (\lambda_{it} l_{it})^{1-\beta} k_{it}^{\beta} \tag{A4}$$

Labor l_{it} is given exogenously, as is labor productivity λ_{it}, which grows at a fixed rate α: $\lambda_{it} = \exp\{\alpha t\}$. Capital k_{it} accumulates with investments i_{it}, assuming zero depreciation.

$$\frac{\mathrm{d}}{\mathrm{d}t} k_{it} = i_{it} \tag{A5}$$

Emissions and Emission Allowances

Greenhouse gas emissions e_{it} are a byproduct of economic activity y_{it}. We assume that the emission intensity falls exogenously due to technological progress at rate ν. Beyond this, emissions may be reduced by investments b_{it} into abatement a_{it}, bringing down the instantaneous emission intensity σ_{it}. Parameters ξ_i describes the effectiveness of these investments, and γ the effectiveness of the abatement option.

$$e_{it} = \sigma_{it} e^{-\nu t} y_{it} \tag{A6}$$

$$\sigma_{it} = (1 + a_{it})^{-\gamma} \tag{A7}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} a_{it} = \xi_i b_{it} \tag{A8}$$

Emissions can exceed allowances q_{it}, which in our model are chosen endogenously by individual regions. Emission allowances may be traded internationally (z_{it} denotes allowance exports by region i), but we exclude intertemporal banking and borrowing, i.e. total imported and exported allowances must be balanced in every period.

$$e_{it} = q_{it} - z_{it} \tag{A9}$$

$$\sum_j z_{jt} = 0 \qquad t = 1, \ldots \tag{A10}$$

Climate Dynamics

Global warming is driven by total global emissions of CO_2 into the atmosphere, which are equal to cumulative total emission allowances $\sum_i q_{it}$. For details on the following climate equations, see Petschel-Held *et al.* (1999).

$$\frac{\mathrm{d}}{\mathrm{d}t} C_t = \zeta \sum_j q_{jt} - \kappa (C_t - C_0) + \psi E_t \tag{A11}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} E_t = \sum_j q_{jt} \tag{A12}$$

Equation (A11) translates global emissions into carbon concentration in the atmosphere C_t. Concentration C_t rises with global allowances (like emissions do), where ζ converts emissions into changes in concentration, and it decreases with the carbon uptake of oceans proportional (with factor κ) to the increase above the pre-industrial level C_0. The final term limits the ocean carbon uptake (to the fraction $1 - \psi/\zeta_\kappa$ in equilibrium).

$$\frac{\mathrm{d}}{\mathrm{d}t} T_t = \mu \log(C_t / C_0) - \phi (T_t - T_0) \tag{A13}$$

Equation (A13) transforms concentration levels into a global mean atmospheric temperature increase T. Parameter μ controls the strength of the temperature reaction due to a change in concentration, whereas parameter ϕ is related to its timing. Together, they can be interpreted as 'climate sensitivity' ($\mu/\phi \cdot log2$), i.e. the equilibrium temperature increase due to a doubling of concentration. In view of the inertia of the climate system, we run the model for 250 years in steps of 10 years.

The climate change damage function Ω_{it} is taken from Nordhaus and Yang (1996):

$$\Omega_{it} = 1 / (1 + \theta_{1i} (T_t)^{\theta_{2i}}) \tag{A14}$$

Parameters θ_{1i} and θ_{2i} describe the vulnerability of region i.

Two sets of 'book keeping' equations complete the model: the budget constraints for consumption and investments for each region at every point in time, as well as the intertemporal budget constraints ensuring that, over the entire time horizon, the import value must equal the export value in each region.

$$y_{it} + m_{it} = c_{it} + i_{it} + b_{it} + x_{it} \tag{A15}$$

$$\int_0^\infty p_t m_{it} \, \mathrm{d}t = \int_0^\infty p_t x_{it} + p_t^z z_{it} \, \mathrm{d}t \tag{A16}$$

Variables m_{it} and x_{it} are imports and exports of region *i*, respectively, and p_t and p_t^z are the prices of goods and allowances respectively.

Solving the Model for the Game's Equilibrium

As detailed in the main text, we are considering a two-stage game of coalition formation in which in the first stage, decisions about *membership* in an international environmental agreement (IEA), and in the second stage decision about *emission allowances* are taken by players.

The game is solved numerically by backward induction, i.e. first we compute PANE for all possible coalitions, then we test these coalitions for internal and external stability according to the following criteria:

$$W_i|_S \geq W_i|_{S\setminus\{i\}} \qquad \text{for } i \in S \quad \text{(internal stability)} \tag{A17}$$

$$W_j|_S > W_j|_{S\cup\{j\}} \qquad \text{for } j \notin S \quad \text{(external stability)} \tag{A18}$$

The computation of the PANE in the second stage is complicated by the fact that we are looking at an intertemporal optimization model, featuring an environmental externality as well as international trade at the same time. To the best of our knowledge, there are no out-of-the-box solvers available to solve such a model in primal form. Lessmann *et al.* (2009) suggest an iterative approach based on Negishi's approach (Negishi, 1972). In this paper, we use a modified version of the iterative algorithm, which works as follows.

Negishi's approach searches for the social planner solution that corresponds to a competitive equilibrium by varying the weights δ_i under the assumption of joint welfare maximization:[12]

$$\max_{\{i_{jt},b_{jt},m_{jt},x_{jt},z_{jt}:j=1\ldots N\}} \sum_{i=1}^{N} \delta_i W_i \tag{A19}$$

subject to equations (A1)–(A15).

Since this approach exploits the fundamental theorems of welfare economics, it cannot be applied to an economy with externalities. In principle, this problem can be circumvented by making any external effect on other players exogenous to the model (converting variables into parameters that are adjusted in an iteration).

In our context, externalities are climate change damages caused by aggregate global emissions. In the Nash equilibrium, players will only anticipate the effect that their emissions have on their own economic output, not, however, the effect this has on other players' output. We can mimic this in a social planner solution by giving each player his own perception of the causal link between emissions and global warming. Instead of equation (A11), which describes one trajectory of concentration C_t, we introduce N equations for C_{it}:

$$\frac{\mathrm{d}}{\mathrm{d}t} C_{it} = \zeta\left(q_{it} + \sum_{j\neq i} \overline{q_{jt}}\right) - \kappa(C_t - C_0) + \psi E_t \qquad \forall_{i\notin S} \tag{A20}$$

---

[12] Note that the intertemporal budget constraint equation (A16), which contains the (*a priori* unknown) market clearing prices, is omitted from the model.

$$\frac{\mathrm{d}}{\mathrm{d}t}C_{it} = \zeta\left(\sum_{k\in S} q_{kt} + \sum_{j\notin S} \overline{q_{jt}}\right) - \kappa(C_t - C_0) + \psi E_t \qquad \forall_{i\in S} \tag{A21}$$

where the allowance choices of other players enter as a fixed value (a parameter, indicated by the bar), and are set to the levels of the corresponding variables during the previous iteration (or some initial value). The sum of allowances in equation (A12) needs to be adjusted analogously, and the temperature equation (A13) will consequently have N instances of T_{it}, too. The temperature change T_{it}, anticipated by player i, will then enter in equation (A14) instead of T_t.

The so modified model is then solved in a nested iteration: in the inner iteration, we solve the model for a given vector $\overline{q} = (\overline{q_{it}})$ of allowance choices repeatedly, updating $\overline{q_{it}} = q_{it}$ at the end of each iteration, i.e. we perform a fixed point iteration of the mapping $q = G(q)$ where G is the best response of players to the exogenously given strategy $\overline{q_{it}}$ of the other players. If the inner iteration converges, it converges to a Nash equilibrium in allowance choices. However, the markets for allowances and private goods may not be a competitive equilibrium. This is what the outer iteration achieves.

The outer iteration follows the standard Negishi approach: we adjust the welfare weights δ_i in the joint welfare function (equation (A19)) until the intertemporal budget constraints (equation (A16)) are satisfied. The resulting equilibrium is the desired PANE.

Numerical Verification of the Equilibrium

We verify the resulting 'candidate' PANE equilibrium strategies in emissions and trade numerically by comparing them to the results of the following maximization problems:

$$\forall_i \max_{\{i_{it}, b_{it}, m_{it}, x_{it}, z_{it}\}} W_i \tag{A22}$$

subject to equations (A1)–(A16) and prices p_t, p_t^z.

Deviations of this model from our solution should be only within the order of magnitude of numerical accuracy, which is what we find (not shown). In particular, simultaneous clearance of all international markets confirms the competitive equilibrium in international trade.

Internalization of Damages

In the model and its solution algorithm outlined above, climate change damages that occur within a region (or coalition of regions) are fully internalized. However, a region may not internalize all of the damages occurring on its territory, e.g. due to lack of information, or if a region represents a large number of countries that do not coordinate their actions. In such cases, full internalization of damages of this group would overestimate the abatement taken by this player. This point was already made in Nordhaus and Yang (1996, p. 743). Therefore, they divide the damages perceived by such a representative player by the number of countries represented by her. This is implemented in our model in the following way: the anticipated climate change damages Ω_{it} in equation (A15) are only a fraction $1/n_i$ of the original right-hand-side of the equation. In equation (A4), we add the remaining damages $\bar{\Omega}_{it}$ that were not anticipated, i.e. $(\Omega_{it} + \bar{\Omega}_{it})$ instead of just Ω_{it}. The parameter $\bar{\Omega}_{it}$ then needs to be updated in an iteration to $\bar{\Omega}_{it} = (n_i - 1)\Omega_{it}$. Through this procedure, full damages take effect even though only a fraction is anticipated.

For the calculations in this study, we set $n_i = 1$ for all players but the CDM supplier, where we chose $n_i = 100$ such that only very little abatement action is taken in the business-as-usual scenario.

References

Altamirano-Cabrera, J. and Finus, M. (2006). 'Permit Trading and Stability of International Climate Agreements', *Journal of Applied Economics*, Vol. 9, No. 1, pp. 19–47.

Asheim, G., Froyn, C., Hovi, J. and Menz, F. (2006). 'Regional versus Global Cooperation for Climate Control', *Journal of Environmental Economics and Management*, Vol. 51, No. 1, pp. 93–109.

d'Aspremont, C. and Gabszewicz, J. J. (1986). 'On the Stability of Collusion', in G. Matthewson and J. E. Stiglitz (eds), *New Developments in the Analysis of Market Structures*, New York, Macmillan, pp. 243–264.

Barrett, S. (1994). 'Self-enforcing International Environmental Agreements', *Oxford Economic Papers*, Vol. 46, pp. 878–894.

Barrett, S. (1997). 'The Strategy of Trade Sanctions in International Environmental Agreements', *Resource and Energy Economics*, Vol. 19, pp. 345–361.

Barrett, S. (2002). 'Consensus Treaties', *Journal of Institutional and Theoretical Economics*, Vol. 158, No. 4, pp. 529–547.

Barrett, S. (2006). 'Climate Treaties and "Breakthrough" Technologies', *The American Economic Review*, Vol. 96, No. 2, pp. 22–25.

Bosetti, V., Carraro, C., Galeotti, M., Massetti, E. and Tavoni, M. (2006). 'WITCH: a World Induced Technical Change Hybrid Model', *The Energy Journal*, Special Issue Hybrid Modelling of Energy Environment Policies: Reconciling Bottom-up and Top-down, pp. 13–38.

Botteon, M. and Carraro, C. (1998). 'Strategies for Environmental Negotiations: Issue Linkage with Heterogeneous Countries', in N. Hanley and H. Folmer (eds), *Game Theory and the Global Environmental*, Cheltenham, Edward Elgar Publishing, pp. 180–200.

Carbone, J., Helm, C. and Rutherford, T. (2009). 'The Case for International Emission Trade in the Absence of Cooperative Climate Policy', *Journal of Environmental Economics and Management*, Vol. 58, No. 3, pp. 266–280.

Carraro, C. and Siniscalco, D. (1993). 'Strategies for the International Protection of the Environment', *Journal of Public Economics*, Vol. 52, No. 3, pp. 309–328.

Carraro, C. and Siniscalco, D. (1997). 'R&D Cooperation and the Stability of International Environmental Agreements', in C. Carraro (ed.), *International Environmental Agreements—Strategic Policy Issues*, Cheltenham, UK, Edward Elgar Publishing, pp. 71–95.

Carraro, C., Marchiori, C. and Oreffice, S. (2009). 'Endogenous Minimum Participation in International Environmental Treaties', *Environmental and Resource Economics*, Vol. 42, No. 3, pp. 411–425.

Chander, P. and Tulkens, H. (1995). 'A Core-theoretic Solution for the Design of Cooperative Agreements on Transfrontier Pollution', *International Tax and Public Finance*, Vol. 2, pp. 279–293.

Courtois, P. and Haeringer, G. (2012). 'Environmental Cooperation: Ratifying Second-best Agreements', *Public Choice*, Vol. 151, pp. 565–584.

Criqui, P., Mima, S. and Viguier, L. (1999). 'Marginal Abatement Costs of CO_2 Emission Reductions, Geographical Flexibility and Concrete Ceilings: an Assessment Using the POLES Model', *Energy Policy*, Vol. 27, pp. 585–601.

Eichner, T. and Pethig, R. (2013a). 'Self-enforcing Environmental Agreements and International Trade', *Journal of Public Economics*, Vol. 102, pp. 37–50.

Eichner, T. and Pethig, R. (2013b). 'Trade Policy and Self-enforcing Environmental Agreements', Presented at the EAERE2013 conference, Toulouse.

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*, Vol. 19, No. 3, pp. 323–358.

Fankhauser, S. and Hepburn, C. (2010). 'Designing Carbon Markets, Part II: Carbon Markets in Space', *Energy Policy*, Vol. 38, No. 8, pp. 4381–4387.

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—a Survey of Current Issues*, Cheltenham, UK and Northamption MA, USA, Edward Elgar Publishing, pp. 82–158.

Finus, M. (2008). 'Game Theoretic Research on the Design of International Environmental Agreements: Insights, Critical Remarks, and Future Challenges', *International Review of Environmental and Resource Economics*, Vol. 2, pp. 29–67.

Finus, M. and Maus, S. (2008). 'Modesty may Pay', *Journal of Public Economic Theory*, Vol. 10, No. 5, pp. 801–826.

Finus, M. and Rundshagen, B. (2003). 'Endogenous Coalition Formation in Global Pollution Control: a Partition Function Approach', in C. Carraro (ed.), *The Endogenous Formation of Economic Coalitions*, Cheltenham, UK, Edward Elgar Publishing, pp. 199–243.

Folmer, H., Mouche, P. and Ragland, S. (1993). 'Interconnected Games and International Environmental Problems', *Environmental and Resource Economics*, Vol. 3, No. 4, pp. 313–335.

Helm, C. (2003). 'International Emissions Trading with Endogenous Allowance Choices', *Journal of Public Economics*, Vol. 87, No. 12, pp. 2737–2747.

Hoel, M. (1992). 'International Environment Conventions: the Case of Uniform Reductions of Emissions', *Environmental and Resource Economics*, Vol. 2, No. 2, pp. 141–159.

Hoel, M. and Schneider, K. (1997). 'Incentives to Participate in an International Environmental Agreement', *Environmental and Resource Economics*, Vol. 9, No. 2, pp. 153–170.

Ichiishi, T. (1981). 'A Social Coalitional Equilibrium Existence Lemma', *Econometrica: Journal of the Econometric Society*, Vol. 49, No. 2, pp. 369–377.

Kim, Y.-G. and Baumert, K. A. (2002). 'Reducing Uncertainty through Dual-intensity Targets', in K. A. Baumert, O. Blanchard, S. Llosa, and J. F. Perkaus (ed.), *Building on the Kyoto Protocol: Options for Protecting the Climate*, Washington, DC, World Resources Institute, pp. 109–133.

Leimbach, M., Bauer, N., Baumstark, L. and Edenhofer, O. (2010). 'Mitigation Costs in a Globalized World: Climate Policy Analysis with REMIND-R', *Environmental Modeling and Assessment*, Vol. 15, No. 3, pp. 155–173.

Lessmann, K. and Edenhofer, O. (2011). 'Research Cooperation and International Standards in a Model of Coalition Stability', *Resource and Energy Economics*, Vol. 33, No. 1, pp. 36–54. doi:10.1016/j.reseneeco.2010.01.001

Lessmann, K., Marschinski, R. and Edenhofer, O. (2009). 'The Effects of Tariffs on Coalition Formation in a Dynamic Global Warming Game', *Economic Modelling*, Vol. 26, No. 3, pp. 641–649.

Manne, A. and Stephan, G. (2005). 'Global Climate Change and the Equity–Efficiency Puzzle', *Energy*, Vol. 30, pp. 2525–2536.

Meckling, J. and Chung, G. (2009). 'Sectoral Approaches for a Post-2012 Climate Regime: a Taxonomy', *Climate Policy*, Vol. 9, No. 6, pp. 652–668.

Negishi, T. (1972). *General Equilibrium Theory and International Trade*, Amsterdam, North-Holland Pub. Co.

Nordhaus, W. and Boyer, J. (2000). *Warming the World: Economic Models of Global Warming*, Cambridge, MA, MIT Press.

Nordhaus, W. D. (2010). 'Economic Aspects of Global Warming in a Post-Copenhagen Environment', Proceedings of the National Academy of Sciences, 107(26), 11721–11726.

Nordhaus, W. D. and Yang, Z. (1996). 'A Regional Dynamic General-equilibrium Model of Alternative Climate-change Strategies', *The American Economic Review*, Vol. 86, No. 4, pp. 741–765.

Petschel-Held, G., Schellnhuber, H.-J., Bruckner, T., Tóth, F. L. and Hasselmann, K. (1999). 'The Tolerable Windows Approach: Theoretical and Methodological Foundations', *Climatic Change*, Vol. 41, No. 3, pp. 303–331.

Philibert, C. (2000). 'How Could Emissions Trading Benefit Developing Countries', *Energy Policy*, Vol. 28, No. 13, pp. 947–956.

Rubio, S. and Ulph, A. (2007). 'An Infinite-horizon Model of Dynamic Membership of International Environmental Agreements', *Journal of Environmental Economics and Management*, Vol. 54, No. 3, pp. 296–310.

Schmalensee, R. (1998). 'Greenhouse Policy Architecture and Institutions', in W. D. Nordhaus (ed.), *Economics and Policy Issues in Climate Change*, Washington, DC, Resources for the Future, pp. 137–158.

Ulph, A. (2004). 'Stable International Environmental Agreements with a Stock Pollutant, Uncertainty and Learning', *The Journal of Risk and Uncertainty*, Vol. 29, No. 1, pp. 53–73.

UNFCCC (2002). 'Report to the Conference of the Parties on its Seventh Session, Held at Marrakesh from 29 October to 10 November 2001 (Marrakesh Accords)—Part II: Action Taken by the Conference of the Parties', Available online at http://unfccc.int/resource/docs/cop7/13a02.pdf

Weikard, H.-P., Wangler, L. and Freytag, A. (2009). 'Minimum Participation Rules with Heterogeneous Countries', *Jena Economic Research Papers 2009-077*.