# International environmental agreements with asymmetric countries: climate clubs vs. global cooperation

Achim Hagen [*,†], Klaus Eisenack [*]

**Abstract**

We investigate whether global cooperation for emission abatement can be improved if asymmetric countries can sign different parallel environmental agreements. The analysis assumes a two-stage game theoretical model. Conditions for self-enforcing sets of agreements and the resulting total emission abatement are determined. After solving the model for the case of constant marginal benefits from abatement and constant marginal costs of abatement with two agreements we relax our assumptions and allow for multiple coalitions with multiple types of asymmetric countries. We then analyze the effect of multiple coalitions for the case of increasing marginal costs of abatement as well as for decreasing marginal benefits of abatement more generally. The results are sensitive to the assumptions on the benefits from abatement. For constant marginal benefits, the possibility of multiple agreements increases the number of cooperating countries and total abatement (compared to the standard case with a single agreement). For decreasing marginal benefits, total emissions are independent of the number of admitted agreements. The paper thus contributes to the emerging discussion on the scope and limits of climate clubs.

[*]Department for Economics and Statistics, Carl von Ossietzky University Oldenburg, Germany
[†]achim.hagen@uni-oldenburg.de

# 1 Introduction

Most game theoretical studies on IEAs assume that there is (at most) one self-enforcing coalition that abates emissions (e.g. Barrett, 1994, 2001; McGinty, 2007; Pavlova and de Zeeuw, 2013). Another frequently used assumption in the theoretical literature is that countries are symmetric (e.g. Barrett, 1994; Asheim et al., 2006). In light of the slow progress in international climate negotiations, the idea of climate clubs is getting increasing attention (e.g. Weischer et al., 2012; Ostrom, 2012; Widerberg and Stenson, 2013) while the aim of negotiating one single universal agreement is identified by some authors as one primary obstacle to a global treaty within the United Nations Framework Convention on Climate Change (UNFCCC) (e.g. Stewart et al., 2013; Falkner et al., 2010). Multiple parallel agreements of subsets of nation states might promise more contributions to the global public good. In particular, asymmetric countries may sort into clubs with similar or complementary properties. Existing studies with asymmetric countries (but at most one coalition) have shown that in some cases global cooperation can be improved (Barrett, 2001; Eisenack and Kähler, 2012; Heugues, 2012). To our knowledge, the idea of climate clubs has not been exhaustively analyzed in the IEA literature so far.

Our theoretical paper explores the potential of climate clubs by assuming different types of countries and allowing for disjoint IEAs. Each country of either type can choose whether to join an agreement or to sign none of them. Each agreement is framed as a (stable or unstable) coalition, and its members act cooperatively. The Nash game between the coalitions and the non-signatories is non-cooperative.

Our research thus extends the seminal work of Barrett (1994) and Carraro and Siniscalco (1993) within the latter stream in that it analyses the internal and external stability of coaltions (D'Aspremont et al., 1983) in a setting with simultaneous play. There is no Stackelberg leadership of one or the other coalition.

We only know of a small literature that investigates the case of multiple IEAs which is complemented by publications focusing on the analysis of IEAs with assymetric countries. Finus and Rundshagen (2001) analyze a model with symmetric countries that play coalition formation games. In equilibrium, their results show that countries form multiple coalition structures. Asheim et al. (2006) also analyze symmetric countries but in an infinitely repeated game. They conclude that for two coexisting agreements a larger number of cooperating signatories can be sustained, compared to the standard case of a single IEA. Finus (2008) and Osmani and Tol (2010) in contrast use simulations to analyze asymmetric countries in more than one coalition and find that with such coalition structures the situation can be improved compared to the standard case with one coalition.

The following paper starts from a different setting and derives analytical results for different classes of cost and benefit functions (constant and decreasing marginal benefits of abatement as well as constant and increasing marginal costs of abatement). In that we are more general than previous work. By choosing a Nash setting we avoid the tideous question about which of the coalitions moves first. In doing so, we confirm but also qualify some of the results from existing studies with different assumptions. Inter alia, we find that for constant marginal benefits, total abatement increases with the number of coalitions, while it remains identical to the standard case for decreasing marginal benefits.

The next section provides a short review of game theoretical literature on IEAs that contributes to the analysis of multiple IEAs with asymmetric countries. Section 3 is devoted to the case of linear costs and benefits of abatement. We first derive the standard case as a benchmark. Then, for two coalitions, we first solve the second stage abatement game, and subsequently the first stage coalition game. The subsequent Section 4 analyzes the effect of multiple coalitions for the case of increasing marginal costs for abatement as well as for decreasing marginal benefits of abatement in a generalized way. Here we focus on the analysis of the second stage of the game, as the decisions in this stage already reveal the effects of multiple parallel climate clubs. A summary and discussion concludes the paper.

## 2 Contributions from the literature

The field of international environmental agreements is treated by a broad economic literature. From the 1990s on, the theoretical literature started to analyze the logic of the formation of coalitions, regarding the environmental game between countries not only as a prisoners dilemma that leads inevitably to the tragedy of common property goods. Various game theoretical models have been developed that analyze IEAs as cooperative and non-cooperative games. Broad overviews of the literature on coalition formation are given by Finus (2001) and Bloch (1997). This section provides a literature review that does not claim to be exhaustive, but rather concentrates on work that directly or indirectly contributes to the analysis of multiple IEAs with asymmetric countries in a non-cooperative game theoretical setting. One strand of the non-cooperative approaches uses reduced-stage game models and depicts coalition formation as a two-stage game with countries deciding in the first stage about joining a coalition before deciding about their emissions the second stage. These models frequently assume that countries are identical and only one single agreement can be signed.

Barrett (2001) modifies these assumptions and allows for asymmetric countries. He uses a simple model with two types of countries that have a binary choice to either abate or pollute. In his model countries have a linear payoff function that depends on the abatement decisions of all countries. Barrett shows that if countries are strongly asymmetric side payments may increase participation in an IEA.

More recently, a number of studies have extended the model of Barrett (2001), also analyzing IEAs with asymmetric countries. McGinty (2007) uses a numerically solved model with 20 asymmetric countries that have convex abatement cost functions and concave benefits from abatement. He finds that with asymmetric countries and transfer payments, IEAs can achieve substantial emissions reductions even when the gains to the IEA are large. Also Biancardi and Villani (2010) and Ruis and de Zeeuw (2010) analyze IEAs, using models that allow for asymmetries between countries and, as the earlier paper of McGinty (2007), rely on numerical exercises.

Fuentes-Albero and Rubio (2010) use a model with asymmetric countries that differ either in their non-linear abatement costs or in environmental damages. The model is solved analytically and shows that transfer payments can improve the level of cooperation especially if countries differ in environmental damages from emissions. Further, the case of two-sided asymmetries in individual quadratic benefits from emissions as well as in individual linear damages is considered by Pavlova and de Zeeuw (2013). They derive analytical results and find that with asymmetries in both cases, large stable coalitions with countries that contribute only little are possible even without transfers but reduce less than small coalitions of countries that contribute substantially. With transfers, also large heterogeneous coalitions perform better.

Finus and Rundshagen (2001) allow for several coalitions with symmetric countries and use reduced stage game concepts that they call coalition formation games. They compare equilibrium coalition structures in different coalition formation games. Their findings of multiple equilibrium coalition structures in different coalition formation games for symmetric countries let them assume that in the case of heterogeneous countries the possibility to form multiple coalitions could lead to better results concerning global emissions as well as global welfare.

Finus (2008) compares different membership models for IEAs using simulations of a model that includes an empirical climate model with twelve regions as well as a game theoretical model for computing stable coalitions. He finds that with heterogenous countries coalition structures can only be stable if countries of similar cost-benefit-structures form coalitions. He also concludes that allowing for separate agreements among countries with similar interests could improve the results of negotiations for IEAs.

Osmani and Tol (2010) formulate the case of two self-enforcing IEAs and additionally consider two asymmetric country types. They assume a three-stage sequence of play of the coalitions and the non-signatories. Their paper mostly focuses on procedures to compute stable coalitions numerically. By computing some numerical examples they show that the possibility of two coalitions could both increase and decrease emission abatement compared to the standard case with one coalition.

Asheim et al. (2006) model the case of symmetric countries and two agreements. The countries are partitioned in two regions, and can chose whether they sign an agreement for that region or not. Marginal benefits of abatement are constant. The model is solved as an infinitely repeated game under different institutional assumptions and renegotiation-proof agreements are identified. For two coexisting agreements, a larger number of cooperating signatories can be sustained, compared to the standard case of a single IEA.

The foregoing literature review shows that there are already different game theoretical approaches contributing to the analysis of multiple IEAs with asymmetric countries. But at the same time it showed clearly that there is still need for further analysis of this field as the existing literature is not generally conclusive. The model developed in the subsequent sections is a step towards closing this gap by deriving analytical results, comparing the cases of one single and multiple parallel IEAs in a more general way than previous work.

# 3 Climate clubs with linear costs and benefits of abatement

For exposition, we start with the most simple formulation of abatement costs and benefits. We consider the case of two country types, and compare the standard setting that allows at most one agreement with a new setting where two parallel agreements are in place. We assume the standard two-stage game structure (see Carraro and Siniscalco, 1993) with countries choosing first to be a signatory or non-signatory of an IEA. In the second stage the signatories choose cooperatively between playing pollute or abate. This choice set is discrete since the model's linear structure excludes interior solutions for abatement. The model assumptions for the case of one agreement follow Barrett (2001), extended by allowing for two parallel agreements.

There are $N$ countries with $N_1$ type 1 and $N_2$ type 2 countries. If a type $i$ country $(i = 1, 2)$ plays abate, it gets the payoff

$$\Pi_i^A = -c + \alpha_i(z_1 + z_2), \tag{1}$$

with the number of countries of type $i$ that play abate denoted by $z_i$. Countries that play pollute get the payoff

$$\Pi_i^P = \alpha_i(z_1 + z_2). \tag{2}$$

The additional benefits from one more type 1 country playing abate are equal to the benefits from one more type 2 country playing abate. This refers to the case of a global public good, e.g. a pollutant that has a global impact no matter where it is emitted, as is the case for greenhouse gases. The asymmetry of the countries is expressed by the parameter $\alpha_i$ where $\alpha_2$ is normalized to $\alpha_2 = 1$ and $\alpha_1 \in [0, 1]$. A type 2 country therefore benefits at least as much as a type 1 country from abatement. It is assumed that the abatement costs $c > 1$, and that the net benefit of the own abatement of each country $i$, $-c + \alpha_i$, is therefore negative. Thus, playing pollute is the dominant strategy if there is no IEA. The Nash equilibrium of a non-cooperative emission game would consequently be unique with all countries playing pollute.

## 3.1 The standard case with one agreement

The standard case is recalled here to introduce the basic notation and to ease comparison with the case of two parallel agreements as detailed in the following subsection. As usual, we proceed by backward induction, solving the second stage of the game first. In the case of one agreement with $k_1$ type 1 signatories and $k_2$ type 2 signatories, the non-signatories all play pollute as a dominant strategy. The aggregate payoff of the signatories is

$$\Pi^S = -cz_1 + \alpha_1 k_1(z_1 + z_2) - cz_2 + k_2(z_1 + z_2). \tag{3}$$

The signatories maximize their payoff $\Pi^S$ cooperatively with respect to $z_i$. The linear payoff function implies the corner solution $z_i^* = k_i$ if

$$\alpha_1 k_1 + k_2 > c. \tag{4}$$

Here and in the following, variables referring to the game equilibrium with one agreement are denoted with a $*$. All signatories of either type play pollute ($z_i^* = 0$) if

$$\alpha_1 k_1 + k_2 < c. \tag{5}$$

To solve the first stage of the game the criteria of internal and external stability (following D'Aspremont et al. (1983)) are applied. In accordance with these, an agreement is stable if no signatory has an incentive to leave the agreement (internal stability) and no non-signatory wants to join the existing agreement (external stability). Formally, an abating coalition is internally stable if

$$\Pi_i^A(k_i) > \Pi_i^P(k_i - 1). \tag{6}$$

As playing pollute is a dominant strategy for non-signatories, this condition is only fulfilled if the signatories choose to abate and would decide to pollute if one signatory would leave the agreement. This leads to 'linchpin' equilibria as the withdrawal of one country would change the decision of all the other signatories from abate to pollute. From condition (4) and (5) we see that $(k_1^*, k_2^*)$ represents a stable and abating coaltion if condition (4) holds and if

$$c > \alpha_1(k_1^* - 1) + k_2^*, \tag{7}$$

$$c > \alpha_1 k_1^* + (k_2^* - 1). \tag{8}$$

Condition (7) implies condition (8) so that internal stability for a single coalition is given if

$$c + \alpha_1 > \alpha_1 k_1^* + k_2^* > c. \tag{9}$$

The criterion of external stability is implied by this condition because playing pollute is a dominant strategy for non-signatories and therefore a non-signatory has no incentive to join an abating and internally stable agreement.

It would be interesting to know conditions where only countries of the same type would sign the same agreement. This can be seen by setting one coalition to size zero in (9). An abating coalition with only type 1 signatories ($z_2^* = k_2 = 0$) is thus possible for

$$c + \alpha_1 > \alpha_1 k_1^* > c, \tag{10}$$

and with only type 2 countries ($z_1^* = k_1 = 0$) if

$$c + 1 > k_2^* > c. \tag{11}$$

Conditions (9) to (11) show that a stable agreement could either consist of both types of countries or of countries of only one type. If all of these three types of stable agreements are possible, the Nash equilibrium of the complete game is not unique.

## 3.2 Abatement decisions and stable coalitions with two agreements

We now assume the possibility of two parallel agreements that take their abatement decisions independently but cooperate internally. Again we proceed by backward-induction, solving the second stage of the game first. Agreement 1 consists of $k_1$ type 1 countries and agreement 2 of $k_2$ type 2 countries. The aggregate payoff of agreement 1 is thus

$$\Pi_1^S = -cz_1 + \alpha_1 k_1(z_1 + z_2). \tag{12}$$

Maximization of $\Pi_1^S$ leads to the corner solutions

$$z_1 = \begin{cases} k_1 \text{ (abate) if } \alpha_1 k_1 > c, \\ 0 \text{ (pollute) if } \alpha_1 k_1 < c. \end{cases} \tag{13}$$

By analogy, the $k_2$ signatories of agreement 2 play

$$z_2 = \begin{cases} k_2 \text{ (Abate) if } k_2 > c, \\ 0 \text{ (Pollute) if } k_2 < c. \end{cases} \tag{14}$$

We see that the decisions of each agreement $i$ depend on the number of its signatories $k_i$ and on the abatement costs $c$, but are mutually independent.

The first stage of the game is now solved by applying the criteria of internal and external stability in analogy to the case of one agreement. As the abatement decisions of the two agreements are mutually independent, the conditions for both agreements to be internally stable can be reduced to (10) and (11). Like in the case of one agreement, the criterion of external stability is always satisfied because playing pollute is a dominant strategy for non-signatories of an abating agreement so that they have no incentive to join an abating agreement.

We now compare the game equilibrium in the standard case with that of two agreements. A set of stable and abating agreements with two agreements is denoted by $(k_1^{**}, k_2^{**})$. By adding (10) and (11), we find that

$$2c + \alpha_1 + 1 > \alpha_1 k_1^{**} + k_2^{**} > 2c, \tag{15}$$

holds. This allows to compare with the case of one agreement. For convenience, we use the notation $K^{**} := \alpha_1 k_1^{**} + k_2^{**}$ to represent a measure for the total abatement by all coalitions. We see from (9) and (15) that $K^{**} > K^*$, so that we can summarize:

**Proposition 1.** *If the marginal benefits and costs of abatement are constant, the total number of abating countries in the case of two agreements is greater than in the case of one agreement.*

Thus, the main conclusion of this section is that global cooperation benefits from parallel agreements if marginal benefits from abatement are constant[1].

# 4 The effects of multiple climate clubs in a generalized setting

Through the following sections we analyze the more general case of $i \in I$ different types of countries and $j \in J$ possible coalitions. One subsection focuses on non-linear benefits, and the other on non-linear costs of abatement. The number of type $i$ countries in coalition $j$ is denoted by $k_i^j$. Countries that do not sign any agreement

[1] We may ask whether countries in a set of two abating coalitions have an incentive to swap their agreement. This question can be answered as follows: If one signatory country of type $i$ would change the agreement, the number of signatories $k_i$ would decrease by one, while the number of signatories in the other agreement would increase. We saw that a decrease in the number of signatories $k_i$ would change the decision of the remaining signatories from abate to pollute. As a consequence, the total number of abating countries would decrease from $k_i^{**} + k_j^{**}$ to $k_j^{**} + 1$. Thus, the profit of every country would decrease. For this reason, no signatory-country has an incentive to change the agreement and thereby reduce its own profit.

may simply be regarded as coalitions of size 1 as their abatement behavior is not relevant for the following argumentation. The abatement decision of each agreement $j$ is characterized by $q^j := (q_1^j, \ldots, q_{|I|}^j)$, where $q_i^j$ is the quantity of abatement for each signatory country of type $i$ in agreement $j$. The global amount of abatement is therefore given by $Q := \sum_{i,j} k_i^j q_i^j$, whereas the total abatement of agreement $j$ is $Q^j := \sum_i k_i^j q_i^j$ and the abatement of all others $Q^{-j} := Q - Q^j$. We will focus on the analysis of the second stage of the game, as the decisions in this stage already reveal the effects of multiple parallel climate clubs.

In the second stage the signatories of each agreement choose their abatement level cooperatively in a simultaneous Nash game between all (given) coalitions and the non-signatories. We denote by $\bar{Q}^{-j}(k_1^j, \ldots, k_{|I|}^j)$ the second stage equilibrium abatement of countries that are not signatories of agreement $j$, and by $\bar{\Pi}_i^j(k_1^j, \ldots, k_{|I|}^j, \bar{Q}^{-j})$ the payoffs of type $i$ countries in coalition $j$. In the first stage of the game, countries choose between joining one of the multiple international agreements or to be a non-signatory by anticipating the effect of their decision of the second stage game equilibrium. Then, an agreement $j$ is internally stable if no signatory country has an incentive to leave the agreement, i.e.

$$\begin{aligned}\forall i, l \neq j: \qquad & \bar{\Pi}_i^j(k_1^j, \ldots, k_i^j, \ldots, k_{|I|}^j, \bar{Q}^{-j}(k_i^j)) \\ & \geq \bar{\Pi}_i^l(k_1^j, \ldots, k_i^j - 1, \ldots, k_{|I|}^j, \bar{Q}^{-j}(k_i^j - 1)), \qquad (16)\end{aligned}$$

and externally stable if no external country has an incentive to join the agreement

$$\begin{aligned}\forall i, l \neq j: \qquad & \bar{\Pi}_i^l(k_1^j, \ldots, k_i^j, \ldots, k_{|I|}^j, \bar{Q}^{-j}(k_i^j)) \\ & \geq \bar{\Pi}_i^j(k_1^j, \ldots, k_i^j + 1, \ldots, k_{|I|}^j, \bar{Q}^{-j}(k_i^j + 1)). \qquad (17)\end{aligned}$$

## 4.1 Climate clubs with increasing marginal costs of abatement

In this section we assume that marginal benefits from abatement are constant and marginal costs of abatement are increasing, such that the payoff for one abating country of type $i$ in agreement $j$ is

$$\Pi_i^j(q_i^j, Q) = \alpha_i Q - c_i(q_i^j), \tag{18}$$

with $c_i(q_i^j)$ being a differentiable, monotonically increasing convex function of the amount of abatement $q_i^j \geq 0$ undertaken by the country, and $\alpha_i > 0$. The aggregated payoff for the signatories of agreement $j$ is therefore given by

$$\Pi^j(q^j, Q) = \sum_i k_i^j \alpha_i Q - \sum_i k_i^j c_i(q_i^j). \tag{19}$$

The following property of global abatement can be deduced from these assumptions. The proof also shows that multiple individually stable climate clubs can coexist. If the countries in these stable coalitions would be forced to join in a single agreement, it would not be stable.

**Proposition 2.** *If the marginal benefits from abatement are linear and the marginal costs of abatement increasing, global emissions abatement increases with the number of individually stable coalitions.*

*Proof.* Each coalition $j$ maximizes $\Pi^j$ with respect to all components of $q^j$ by taking the total abatement of all others $Q^{-j}$ as given. The first order condition for each country type $l$ in a coalition $j$ is

$$\begin{aligned}\frac{d\Pi^j}{dq_l^j} &= \sum_i k_i^j \alpha_i \frac{dQ}{dq_l^j} - \sum_i k_i^j \frac{d}{dq_l^j} c_i(q_i^j) \\ &= \sum_i k_i^j \alpha_i k_l^j - k_l^j c_l'(q_l^j) = 0,\end{aligned} \tag{20}$$

so that

$$\forall l \in I : \qquad q_l^j = c_l'^{-1}\left(\sum_i k_i^j \alpha_i\right). \tag{21}$$

Thus

$$Q^j = \sum_l k_l^j c_l'^{-1}\left(\sum_i k_i^j \alpha_i\right), \tag{22}$$

does not depend on $Q^{-j}$: Every coalition has a dominant strategy, taking its abatement decision independently of the decision of all other coalitions. As $\frac{d\Pi^j}{dq_l^j}$ is independent of $Q^{-j}$ the stability conditions (16) and (17) are not affected by any abatement decisions of non-signatories of agreement $j$. Thus, the stability of coalition $j$ is independent of the existence of other individually stable abating coalitions. The maximum number of countries within stable agreements satisfying (16) and (17) increases with the number of agreements. By this argument, multiple coalitions can include a larger number of cooperatively abating countries in stable agreements than one single IEA. If there exists at least one individually stable abating coalition satisfying (16) and (17) [2] with $\sum_i k_i^j > 1$, and if there are enough countries to form a further individually stable abating coalition, they will do so. Their abatement will be additional to other countries' amount without causing existing coalitions to increase their emissions. □

This case can roughly be summarized as follows. With constant marginal benefits of abatement each coalition makes an independent abatement decision. Some countries may prefer to form a stable coalition (of size larger than one) to reap some benefits of cooperation. We know from the established literature that such stable coalitions tend to be small. If additional countries would join such a coalition, the cooperation benefits would diminish. However, countries which are not part of an

[2] An in depth analysis of individually stable and abating coalitions in the setting that admits for only one coalition may be found, e.g., in Fuentes-Albero and Rubio (2010).

existing coalition may like to be willing to form another small coalition with larger cooperation benefits within. Due to the dominant strategies, these cooperation benefits are not affected by the number of countries that already cooperate in other stable coalitions. Thus, as long as there is at least one set of countries left that would form a stable coalition (even if there would be no other coalition), these countries would cooperate, leading to emissions abatement that is additional to that of other coalitions.

## 4.2 Multiple coalitions with decreasing marginal benefits of abatement

We now assume constant marginal abatement costs and decreasing marginal benefits from abatement. The payoff of an abating country thus depends on the global abatement as well as on its own decision and takes the form

$$\Pi_i^j(q_i^j, Q) = f_i(Q) - \gamma_i q_i^j, \tag{23}$$

with $f_i(Q)$ being a differentiable, monotonically increasing concave function of the global quantity of abatement $Q$, $q_i^j \geq 0$ and $\gamma_i > 0$. The aggregate payoff of a coalition $j$ is given by

$$\Pi^j(q^j, Q) = \sum_i k_i^j f_i(Q) - \sum_i \gamma_i k_i^j q_i^j. \tag{24}$$

We obtain a result under these conditions that is in stark contrast to the case with constant marginal benefits.

**Proposition 3.** *If marginal costs of abatement are constant and marginal benefits of global abatement are decreasing, at most one stable coalition will decide to undertake abatement efforts.*

*Proof.* Each coalition $j \in J$ maximizes $\Pi^j$ with respect to all components $q^j$, taking the abatement of the other countries as given. The derivative of coalition $j$'s payoff by abatement of a member country of type $l$ is

$$\frac{d\Pi^j}{dq_l^j} = \sum_i k_i^j f_i' \frac{dQ}{dq_l^j} - \sum_i \gamma_i k_i^j \frac{dq_i^j}{dq_l^j} \tag{25}$$

$$= \sum_i k_i^j f_i'(Q) k_l^j - \gamma_l k_l^j. \tag{26}$$

The optimal abatement decision is either an interior solution with $\frac{d\Pi^j}{dq_l^j} = 0$, or a corner solution with $q_l^j = 0$. Together with the equation $Q = \sum_{i,j} q_i^j$, the Nash equilibrium is thus characterized by $|I| \cdot |J| + 1$ conditions for the $|I| \cdot |J| + 1$ variables $Q, (q_i^j)_{i \in I, j \in J}$.

We now show by contradiction that there cannot be two abating coalitions. So, suppose there are two different stable and abating coalitions $j, k$. This means that there are at least two country types $l, m$ (possibly $l = m$) with both $q_l^j, q_m^k > 0$, i.e. both abatement levels are interior solutions. It thus holds that

$$\gamma_l = \sum_i k_i^j f_i'(Q), \tag{27}$$

$$\gamma_m = \sum_i k_i^k f_i'(Q). \tag{28}$$

This are two conditions for just one free variable. Thus, except for a boundary case, they cannot hold simultaneously. Thus, the assumption leads to a contradiction. □

The proposition shows that considering multiple climate clubs leads to no improvements under the settings of this section. There is no incentive to create a second coalition, and thus global abatement cannot be increased. If there would be a group of countries that is able to form a stable and abating coalition if it is the only coalition, it would ultimately refrain from cooperation if there is already another abating coalition in place. Any club of countries would freeride if some already abate.

# 5 Conclusions

Our paper has analyzed the effects of allowing for multiple international environmental agreements (IEAs) when there are asymmetric countries of multiple types. In a two-stage game, countries first choose whether they sign one agreement, or to be a non-signatory. In the second stage, each coalition acts as a unitary actor in a non-cooperative Nash game between the coalitions and the non-signatories. We compare emissions abatement and coalition stability in the multiple IEAs case with the standard case where at most one IEA is possible. We investigate this for constant as well as decreasing marginal benefits from abatement and for constant as well as increasing marginal costs of abatement.

For constant marginal benefits, multiple IEAs lead to more total abatement and to a larger number of cooperating countries in multiple "climate clubs". Interestingly, this effect does not depend on the shares of the country types within the set of all countries in the game. These results follow from the dominant abatement strategies of the coalitions. In the special case with marginal benefits from abatement as well as marginal abatement costs being linear, these dominant strategies follow the linchpin character of the game equilibrium. One IEA is self-enforcing if all countries would chose to pollute, supposed one more country is leaving the IEA. This effect is replicated for each IEA. Thus two coalitions are stabilized with more abatement than just one.

When marginal benefits decrease and marginal costs are constant, this picture changes. As there is no equilibrium structure with more than one abating stable coalition in the second game stage, only one agreement will abate emissions cooperatively. All other countries would refrain from cooperation regardless of their potential membership in another coalition. Therefore the possibility of multiple coalitions does not lead to improvements compared to the case with only one agreement.

The comparison of the different cases shows that the effect of climate clubs substantially depends on qualitative properties of abatement benefit functions, even if they are quite simple. In this general sense, our results are in line with the ambiguity results in the examples of Osmani and Tol (2010). In contrast, however, we can generally show for our assumptions that climate clubs are at least not detrimental to global cooperation. It would require further consideration whether the positive effects shown by Asheim et al. (2006) mostly stem from the constant marginal benefits assumption.

Nevertheless, our results need to be taken with precaution. Although our analysis is more general than single numerical examples, it sticks to either linear cost or benefit functions. This requests for further generalisation, including the case of both nonlinear costs and benefits at the same time. Also the welfare effects and the comparative statics require more attention. It would further be interesting to determine intercoalition stability (Osmani and Tol, 2010) for the case of multiple parallel abating agreements more explicitly. On the other hand, the paper already shows how different assumptions lead to different effects of climate clubs. We think that our analysis is thus a consequent stepping stone towards a more detailed understanding of the determinants for beneficial or detrimental effects of climate clubs.

Further analysis would profit from analyzing more carefully how countries would voluntarily sort into parallel agreements. Do countries of the same type join, or would types mix in IEAs? In any case, we need to conclude that the idea that climate clubs do benefit global climate protection has to be taken with precaution, but that it deserves more analytical attention.

# References

Asheim, G. B., C. B. Froyn, J. Hovi, and F. C. Menz (2006). Regional versus global cooperation for climate control. *Journal of Environmental Economics and Management 51*(1), 93–109.

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers 46*, 878–894.

Barrett, S. (2001). International cooperation for sale. *European Economic Review 45*(10), 1835–1850.

Biancardi, M. and G. Villani (2010). International environmental agreements with asymmetric countries. *Computational Economics 36*, 69–92.

Bloch, F. (1997). *New directions in the economic theory of the environment*, Chapter Non-cooperative models of coalition formation in games with spillovers, pp. 311–352. Cambridge University Press.

Carraro, C. and D. Siniscalco (1993). Strategies for the international protection of the environment. *Journal of Public Economics 52*(3), 309–328.

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

Eisenack, K. and L. Kähler (2012). Unilateral emission reductions can lead to pareto improvements when adaptation to damages is possible. Technical Report Oldenburg Discussion Papers in Economics V – 344 – 12.

Falkner, R., H. Stephan, and J. Vogler (2010). International climate policy after copenhagen: Towards a 'building blocks' approach. *Global Policy 1*, 252–262.

Finus, M. (2001). *Game theory and international environmental cooperation*. Edward Elgar.

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 2*, 29–67.

Finus, M. and B. Rundshagen (2001). Endogenous coalition formation in global pollution control. Technical report, Fondazione Eni Enrico Mattei.

Fuentes-Albero, C. and S. J. Rubio (2010). Can international environmental cooperation be bought? *European Journal of Operational Research 202*, 255–264.

Heugues, M. (2012). Endogenous timing in pollution control: Stackelberg versus cournot-nash equilibria. *Strategic Behavior and the Environment 2*(2), 133–158.

McGinty, M. (2007). International environmental agreements among asymmetric nations. *Oxford Economic Papers 59*, 45–62.

Osmani, D. and R. Tol (2010). The case of two self-enforcing international agreements for environmental protection with asymmetric countries. *Computational Economics 36*, 93–119.

Ostrom, E. (2012). Nested externalities and polycentric institutions: must we wait for global solutions to climate change before taking actions at other scales? *Economic Theory 49*(2), 353–369.

Pavlova, Y. and A. de Zeeuw (2013). Asymmetries in international environmental agreements. *Environment and Development Economics 18*, 51–68.

Ruis, A. and A. de Zeeuw (2010). International cooperation to combat climate change. *Public Finance and Management 10*, 379–404.

Stewart, R. B., M. Oppenheimer, and B. Rudyk (2013). A new strategy for global climate protection. *Climatic change 120*, 1–12.

Weischer, L., J. Morgan, and M. Patel (2012). Climate clubs: Can small groups of countries make a big difference in addressing climate change? *Review of European Community and International Environmental Law 21*(3), 177–192.

Widerberg, O. and D. E. Stenson (2013). *Climate clubs and the UNFCCCO - Complement, bypass or conflict?* Fores.