Self-Enforcing Climate Coalitions and Preferential Free Trade Agreements

Thomas Kuhn†
Department of Economics, Chemnitz University of Technology

Radomir Pestow††
Department of Mathematics, Chemnitz University of Technology

Anja Zenker‡
Department of Economics, Chemnitz University of Technology

- Preliminary Version -

Abstract

In this paper we discuss the endogenous formation of self-enforcing climate coalitions when it is being linked to the issue of a preferential free trade agreement. As a framework a strategic trade model is used in which countries may discourage greenhouse gas emissions by means of an import tariff on dirty goods. In addition, countries can set strategic emissions caps being effective on a permits market. Our main focus, however, is on the incentives provided to members of a preferential free trade area when terms of trade gains are utilized for combatting global warming. We propose strong evidence for that the welfare gains of trade liberalization are strongly promoting the formation of climate coalitions. In the parametrical simulation of the model global emissions as well as climate change damages are found significantly reduced compared to the BAU scenario, while global welfare is found significantly improved.

JEL classification: Q54, Q58, F18, Q56, C72

Key words: Climate Change, Self-enforcing Environmental Agreements, Free Trade Area, Strategic Trade Policy

---

† e-mail: thomas.kuhn@wirtschaft.tu-chemnitz.de, ph.: +49 371 531 34941
†† e-mail: radomir.pestow@gmail.com
‡ e-mail: anja.zenker@wirtschaft.tu-chemnitz.de, ph.: +49 371 531 39967

1 Introduction

When it comes to the combat of global warming, broad collective action is called for to achieve an effective mitigation of global greenhouse gas emissions. Multilateral cooperation among countries is typically institutionalized in the form of international environmental agreements (IEAs) such as the Kyoto Protocol. Since there is no international institution delegated by nations to enforce an efficient allocation of emissions from a global perspective and by means of supranational legislation, self-enforcing voluntary agreements are required. They should provide incentives to states for reducing their emissions to a level which is considered consistent with a global temperature rise of about less than 2°C.

Self-enforcement is often associated with the criterion of coalition stability which is defined as a situation in which no signatory state of the climate coalition has an incentive to leave the coalition and no non-signatory state has an incentive to join in. Thus, cooperation on the governance of a global public good such as the mitigation of greenhouse gas emissions can only be sustained if an IEA is set up such that it can prevent countries from taking a free ride. From the early 1990s onwards, a comprehensive body of literature has emerged, focusing on the game-theoretic analysis of self-enforcing IEAs. Seminal papers of the non-cooperative literature include *Hoel* (1992), *Carraro and Siniscalco* (1993), and *Barrett* (1994). Assuming that the game is carried out in a Cournot-Nash-fashion, the results suggest that self-enforcing IEAs can only be implemented if either the number of signatories is small or if the agreement commits to lax reduction targets compared to the business-as-usual emission paths, irrespective of whether climate damage takes a linear or quadratic form (*Hoel,* 1992, *Barrett,* 1994). Put differently, there appears to be a trade-off between the effectiveness of an IEA and its stability (*Finus*, 2003).

In order to pave the way for a global treaty with ambitious emission reduction targets, several treaty mechanisms have been considered. A comprehensive overview of the literature on compensation measures such as side-payments or non-material payoffs is provided by *Finus* (2003). The findings can be summarized as follows: Since transfers must be self-financed, i.e. they must result from the welfare gain of the coalition, they hardly can improve cooperation among symmetric countries (*Carraro and Siniscalco*, 1993, *Barrett*, 1997a). If countries are heterogeneous, transfers may lead to larger stable coalition structures but the results highly depend on the specific design of both, the allocation rules and the coalition formation (*Botteon and Carraro*, 1997, *Barrett*, 1997a, *Eyckmans and Finus*, 2007). According to *Finus* (2003), transfer commitments made before the coalition has formed (ex-ante transfers) turn out to be less effective than transfer commitments made to enlarge an already existing coalition (ex-post transfers), though the term ‘commitment’ must be conceived as in line with the self-enforcement requirement in order to be credible.[1] Overall, the scope of transfers to improve the prospects of a broad and effective cooperation appears to be rather limited (*Eyckmans and Finus*, 2007).

Other instruments to induce cooperation reviewed in the literature involve punishments and sanctions. Even though trade sanctions such as border tax adjustments are particularly popular as effective threats to combat carbon leakage (*Bucher and Schenker*, 2011, *Fischer and Fox*, 2012), there is a dispute about their imposition concerning their credibility on the one hand as well as their compliance with the non-discrimination rules of the *WTO* on the other hand. *Barrett* (1997b) argues that, in principle, the use of trade sanctions to increase the provision of a global public good can be welfare-improving and enhance cooperation if the threat of imposition is credible. However, the credibility of trade sanctions is inherently subject to incentives to weaken sanctions in case they might imply a welfare loss for signatory states, too. Therefore, this instrument is also limited in terms of raising participation in IEAs.

More recently, attention has been paid to the linkage of negotiation issues, or, more precisely, to the linkage of the public-good agreement to a club-good agreement. It is assumed that a simultaneous membership in both agreements is required. Here, the potential for success in terms of participation and stability crucially depends on the linked issue. Most papers focus on R&D cooperation such as *Carraro and Siniscalco* (1997), *Carraro* (1999), and *Kemfert* (2004). Two findings are striking: first, R&D cooperation provides a competitive advantage to signatories to produce at lower unit costs by exploiting a more efficient technology, but, even under most favorable assumptions, that tends to disappear when the number of signatories increases and more countries share the same technology. Hence, the R&D cooperation issue entails diminishing returns with respect to the coalition size, whereby it may be optimal to exclude some countries from the joint cooperation on R&D and climate change mitigation. The implication of non-monotonic payoff functions is that some non-signatories would like to join the coalition but are excluded from doing so (*Carraro*, 1999), violating the condition of external stability. In light of these results, the objective of a large, even full participation, that is, a global climate coalition, might not even be desired by the countries which signed the linked agreement. Second, it is reasonable to assume that R&D cooperation entails spillover effects to non-signatories

---

[1] For this purpose, the notion of internal stability has later been modified to a concept of potential internal stability. It is stated that a coalition is *potentially* internally stable if it can be stabilized by a self-financed transfer scheme such that no signatory state – neither the net payers nor the net recipients – has an incentive to behave as a free rider (*Bosetti et al.*, 2013).

through trade of technological innovations and capital flows used for R&D investment (*Kemfert*, 2004). These spillovers might further reduce the gains from the coupled club good.

This is why trade linkage could be more promising by preserving the excludability of the benefits associated with preferential trade cooperation. Previous approaches by *Barrett* (1997b) and *Finus and Rundshagen* (2000) focused on protectionist trade policies that are implemented vis-à-vis non-signatories. As we have already discussed trade sanctions as a punishment mechanism, we will only refer to the results of an IEA linked to a custom union when plant location is endogenous. In such a case, issue linkage can even reduce participation and global welfare if strategically behaving countries give rise to eco-dumping (*Finus and Rundshagen*, 2000).

In this paper, in contrast to a harmonized trade policy vis-à-vis non-signatories, we employ a model in which climate negotiations are interlinked to negotiations on a preferential free trade area (PFTA) while strategic trade and environmental policies vis-à-vis non-signatories are carried out individually. Thereby, excludable benefits are generated by the preferential trade liberalization only. Furthermore, signatories do not necessarily harmonize their tariff rates, but they are chosen with the objective of maximizing the coalition’s welfare. Hence, incentives to deviate from the agreed policies do not occur. As proposed by *Leal-Arcas* (2013), such interlinked trade-climate agreements could be implemented by including climate-related provisions within preferential trade agreements such as the TTIP.

The model employed in the paper is an extension as well as an appropriate modification of the Stackelberg leader-follower framework by *Eichner and Pethig* (2013a, 2013b). However, in contrast to the existing literature analyzing the impact of traditional trade barriers on climate change, we introduce a free trade area as an incentive device for the formation of climate coalitions. Therefore, we have to identify the trade pattern within the free trade area and strictly distinct it from the trade patterns existing outside the area. This is basically done by a novel differentiation of firms’ supplies with respect to the target markets which also implies an appropriate modification of the equilibrium concept for local markets.

The paper is organized as follows. The next section provides the model with a focus on the microfoundations concerning the market equilibria and the trade patterns. In section 3 the strategic policies of fringe and coalition countries are modelled in a strategic Stackelberg leader-follower framework at a given coalition size. In section four the endogenous formation of coalitions is presented including the discussion on the external and internal stability. In section 5 the results of the numerical simulation of the analytical model are given followed by some concluding remarks in section 6.

2 Microfoundations

In the model, we consider a world economy composed of n countries. Each country $i = 1, \dots, n$ has an endowment $\bar{r}$ of a (composite) production factor at its disposal which can be used for the production of either a 'clean' consumer good, x_i, or a 'dirty' consumer good, e_i. The clean good serves as a numeraire while the dirty good for instance may represent some form of a fossil energy carrier. In each country there is a perfectly competitive firm serving the domestic as well as potentially each foreign market. The supply of the dirty good is differentiated according to the country of destination because firms not only must cover the transportation cost at least to some extent but as well must meet country-specific regulations and standards. Therefore, it is reasonable to assume that the opportunity cost of the dirty good may vary dependent on the respective market on target.

A firm in country i produces a dirty good using a decreasing returns-to-scale technology which is decomposed into a number of n specific supplies according to the target markets. For the clean good a constant returns to scale technology is used. Formally, let e_{ij} , $i, j = 1, \dots, n$ denote the amount of the dirty good produced in country i to be shipped to country j such that the first index represents the country of origin and the second one represents the country of destination. Further, let α_x, $\alpha_{e_{ij}}$ give the technology coefficients for the clean and decomposed dirty good, respectively, and let r_x, $r_{e_{i1}}$ give the respective factor inputs. Then the country's factor constraint takes the form:

$$\bar{r} = r_x + r_{e_{i1}} + \cdots + r_{e_{in}}. \tag{1}$$

Furthermore, the production functions for the clean and the decomposed dirty good are defined as follows:

$$\begin{aligned} x_i &= \alpha_x r_x, && i = 1, \dots, n \\ e_{ij} &= \sqrt{\frac{r_{e_{ij}}}{\alpha_{e_{ij}}}}, && i, j = 1, \dots, n. \end{aligned} \tag{2}$$

For the sake of simplicity, we make use of the following specific technology coefficients:

$$\alpha_x = 1,$$
$$\alpha_{e_{ij}} = \begin{cases} \alpha_H & if\ i = j \\ \alpha^* & if\ i \neq j \end{cases} \qquad i,j = 1,\dots,n. \quad (3)$$

In case of the dirty good the assumption made is $0 < \alpha_H < \alpha^*$, that is, opportunity costs shall not differ among the various foreign destinations but are generally considered being higher for cross-border trade due to greater transportation and administrative efforts.

From there, taking into account the maximum producible amount of the clean good given by $\bar{x} = \alpha_x \bar{r}$, we can derive the quadratic production possibility frontier[2] of country i:

$$x_i^S = T\left(e_{i1}^S, \dots, e_{in}^S\right) = \bar{x} - \alpha_x \sum_{j=1}^n \alpha_{e_{ij}} \left(e_{ij}^S\right)^2 = \bar{x} - \left(\alpha_H \left(e_{ii}^S\right)^2 + \alpha^* \sum_{\substack{j=1, \\ j \neq i}}^n \left(e_{ij}^S\right)^2\right), \qquad (4)$$

where T is found a decreasing and strictly concave function in any e_{ij}^S.

With respect to the demand side, we adhere to the framework by *Eichner and Pethig* (2013a, 2013b) and take a representative consumer in country i to maximize a quasi linear utility function[3]

$$U_i\left(x_i^D, e_1^D, \dots, e_n^D\right) = V_i\left(e_i^D\right) + x_i^D. \qquad (5)$$

The marginal utility of the dirty good $V(.)$ is positive, but decreasing in the dirty good e_i^D, whereas it is constant in the numeraire good x_i^D as usual. Moreover, consumers do not discriminate between domestic supplies and imports since commodities in their view are homogeneous. In the following we use the specific form

$$V_i\left(e_i^D\right) = a e_i^D - \frac{b}{2}\left(e_i^D\right)^2, \qquad (6)$$

with parameters a, b being positive.

The reason why e represents a dirty good simply lies in the fact that it is coupled with greenhouse gas emissions like carbon dioxide which naturally are modelled as a global public externality. Hence, the damage function takes the usual form

$$D\left(\sum_{j=1}^n e_j^D\right) \qquad (7)$$

which, later on, will show up in the welfare function of any country. Here, the basic assumption is that the consumption of the dirty good is generating an emission one for one

---

[2] Please note that the superscript S indicates quantities supplied.
[3] Please note that the superscript D indicates quantities demanded.

and thus global emission is given by the sum $\sum_{j=1}^{n} e_j^D$. Marginal damages are increasing and the following specification is adopted:

$$D\left(\sum_{j=1}^{n} e_j^D\right) = \frac{\delta}{2}\left(\sum_{j=1}^{n} e_j^D\right)^2, \tag{8}$$

with parameter $\delta > 0$.

Global damages affect the welfare of any single country and cannot be ignored apart from any free-riding incentives. However, countries which opt for free riding may view their impact on global warming negligible compared to the cost of emission abatement. This is exactly the challenge in the combat of a global public bad like it is the case with global warming. The more important is the formation of climate coalitions.

National governments in principle have two kinds of policy instruments available: a national system of emission permits trading and a trade tariff which in case does not only simply works in the traditional way but also may address environmental disruptions. In order to reduce carbon emissions by means of emissions permits, each government is able to set a national cap $e_i > 0$ and is auctioning the number of available emission permits e_i at a permit price, π_i. The households are those agents which are required to hold a permit for consuming the dirty good one for one to internalize, more or less perfectly, the externality. Additionally, governments can impose a trade tariff $t_i \in \mathbb{R}$ whose algebraic sign is unconstrained, i.e. it may take the form of an import tariff ($t_i > 0$) or of an export tax ($t_i < 0$). Put differently, an import tariff $t_i > 0$ combines a tax on fuel consumption with a subsidy on local fuel production to the advantage of domestic firms, while an export tax $t_i < 0$ combines a subsidy on fuel consumption with a tax on fuel production to the disadvantage of domestic firms.

The tariff design is equivalent to a unit tax that decouples the domestic consumer price of the dirty good from the foreign producer price such that the domestic producer face the domestic consumer price $p_i + t_i$, $t_i \in \mathbb{R}$ whereas foreign producers only receive the price p_i net of the import tariff to be paid.

However, this arrangement gets considerably modified if one intends to establish a free trade area. In this case, policies of countries that are part of the free trade area discriminate between trade with coalition member states and that with non-member states. This is equally true for a climate coalition which is building upon a free trade area as an incentive to combat global warming as will become obvious below. As a result, producers from coalition and fringe countries face different prices prevailing for their exports to local markets in coalition countries. Naturally, the free trade arrangement privileges the firms of its member countries.

Formally, countries first need to be sorted according to their group membership. If country i is a member of the climate and free trade coalition $C := \{1,2,\dots,m\}$, it will be called a coalition country $i \in C$. Accordingly, we name country i a fringe country if it is not a coalition member, i.e. $i \notin C$, that is $\in F := \{m+1, m+2, \dots, n\}$, $m \leq n$. Now, the tariff design must take into account that producers from coalition countries, $j \in C$, are generally exempted from an import tariff imposed by any other coalition country, $i \in C$, $i \neq j$, in accordance with the free trade arrangement. Hence, these firms receive the consumer price $p_i + t_i$, $t_i \in \mathbb{R}$ while producers from fringe countries, $j \in F$ just receive the producer price p_i.

Now, taking the local market prices $p_1, \dots, p_n$ and the tariff rates $t_1, \dots, t_n$ as given, a firm in country i maximizes its profits subject to the production possibility frontier in (4) by optimally choosing the supplies of the clean as well as the differentiated dirty commodity, $x_i^S, e_{i1}^S, \dots, e_{in}^S$. The profit functions for firms are therefore given as follows:

$$\pi_i\left(x_i^S, e_{i1}^S, \dots, e_{in}^S\right) = x_i^S + (p_i + t_i)e_{ii}^S + \sum_{\substack{j=1,\\ j\in C,\\ j\neq i}}^{m} \left(p_j + t_j\right)e_{ij}^S + \sum_{\substack{j=m+1,\\ j\notin C}}^{n} p_j e_{ij}^S, \qquad \text{for } i \in C \qquad (9a)$$

$$\pi_i\left(x_i^S, e_{i1}^S, \dots, e_{in}^S\right) = x_i^S + (p_i + t_i)e_{ii}^S + \sum_{\substack{j=1,\\ j\neq i}}^{n} p_j e_{ij}^S, \qquad \text{for } i \notin C. \qquad (9b)$$

The optimal outputs derived from the first-order conditions of the profit maximization problem yield:

$$e_{ii}^S = \frac{p_i + t_i}{2\alpha_H}, \qquad \left(e_{ij}^S\right)_{\substack{j\in C,\\ j\neq i}} = \frac{p_j + t_j}{2\alpha^*}, \qquad \left(e_{ij}^S\right)_{j\notin C} = \frac{p_j}{2\alpha^*}, \qquad \text{for } i \in C \qquad (10a)$$

$$e_{ii}^S = \frac{p_i + t_i}{2\alpha_H}, \qquad \left(e_{ij}^S\right)_{j\neq i} = \frac{p_j}{2\alpha^*}, \qquad \text{for } i \notin C. \qquad (10b)$$

Turning to the demand side, a representative consumer in a coalition or fringe country i is naturally facing the tax-inclusive price of the dirty good, $p_i + t_i$, charged by all suppliers irrespective of their origin. Since global emissions and their impact on climate change are external in the consumers' view and thus are not taken into account, the demands for the clean and dirty good, x_i^D and e_i^D, respectively, are chosen to maximize utility, $U_i(.)$, subject to the budget constraint

$$y_i = x_i^D + (p_i + t_i + \pi_i)e_i^D. \qquad (11)$$

The income of a representative consumer in country i, denoted by y_i, comprises producer rents, permit income, as well as tariff income, because of the instant transfer of all kinds of income generated in the economy back to the consumer. According to this definition, the income functions read

$$y_i := x_i^S + (p_i + t_i)e_{ii}^S + \sum_{\substack{j=1,\\ j\in C,\\ j\neq i}}^{m} (p_j + t_j)e_{ij}^S + \sum_{\substack{j=m+1,\\ j\notin C}}^{n} p_j e_{ij}^S + \pi_i e_i^D + t_i \sum_{\substack{j=m+1,\\ j\notin C}}^{n} e_{ji}^S, \quad \text{for } i \in C \quad (12a)$$

$$y_i := x_i^S + (p_i + t_i)e_{ii}^S + \sum_{\substack{j=1,\\ j\neq i}}^{n} p_j e_{ij}^S + \pi_i e_i^D + t_i \sum_{\substack{j=1,\\ j\neq i}}^{n} e_{ji}^S, \quad \text{for } i \notin C \quad (12b)$$

This particular income will later on get determined along with the market equilibria. The consumer, however, takes income as given.

The demand for the dirty good arises from the first-order conditions of the utility maximization problem:

$$e_i^D(p_i, t_i, \pi_i) = \frac{a - (p_i + t_i + \pi_i)}{b}. \qquad (13)$$

As the demand for the dirty good also invokes an equal demand for permits in the national emissions trading scheme, the local permit market is in equilibrium if the following condition holds:

$$e_i^D(p_i, t_i, \pi_i) = e_i, \qquad (14)$$

with e_i being the emission cap set by the national government. This yields the equilibrium permit price π_i^*.

Furthermore, there is a world market for good X sold at world price set to $p_x \equiv 1$, and, moreover, there is a local market for the dirty good sold at local price p_i in each country i. These markets are in equilibrium if the following conditions hold simultaneously:

$$\sum_{j=1}^{n} x_j^D = \sum_{j=1}^{n} x_j^S, \qquad e_i^D = \sum_{j=1}^{n} e_{ji}^S, \qquad i = 1, \dots, n. \qquad (15)$$

One should note that the total supply to the local market in country i originates from the aggregated exports of all $j \neq i$ firms into country i in addition to the supply of the domestic firm, and can be itemized as:

$$\sum_{j=1}^{n} e_{ji}^S = e_{ii}^S + \sum_{\substack{j=1,\\ j\in C,\\ j\neq i}}^{m} e_{ji}^S + \sum_{\substack{j=m+1,\\ j\notin C,\\ j\neq i}}^{n} e_{ji}^S. \qquad (16)$$

The equilibrium condition can be further specified by substituting the supplied quantities found in (10) for the RHS in (16), and by substituting demand from (13) for the LHS in (15) along with the emission cap from (14). Then the equilibrium conditions for the local markets in coalition as well as in fringe countries read:

$$e_i = \frac{p_i + t_i}{2\alpha_H} + (m-1)\left(\frac{p_i + t_i}{2\alpha^*}\right) + (n-m)\left(\frac{p_i}{2\alpha^*}\right), \qquad \text{for } i \in C \qquad (17a)$$

$$e_i = \frac{p_i + t_i}{2\alpha_H} + (n-1)\left(\frac{p_i}{2\alpha^*}\right), \qquad \text{for } i \notin C. \qquad (17b)$$

The resulting local equilibrium prices of the dirty good are

$$p_i^*(e_i, t_i) = \frac{2\alpha_H\alpha^* e_i - \alpha^* t_i - (m-1)\alpha_H t_i}{(n-1)\alpha_H + \alpha^*}, \qquad \text{for } i \in C \tag{18a}$$

$$p_i^*(e_i, t_i) = \frac{2\alpha_H\alpha^* e_i - \alpha^* t_i}{(n-1)\alpha_H + \alpha^*}, \qquad \text{for } i \notin C. \tag{18b}$$

As can be seen from (18a) and (18b), the government of country i can impact the local price of the dirty good by adapting its environmental and trade policies.

The equilibrium output of the dirty good *generated* in country $i = 1, \dots, n$, (and in part exported) can be determined by making use of the overall supply function of the representative firm, e_i^S:

$$e_i^S = e_{ii}^S + \sum_{\substack{j=1,\\ j\in C,\\ j\neq i}}^{m} e_{ij}^S + \sum_{\substack{j=m+1,\\ j\notin C}}^{n} e_{ij}^S, \qquad \text{for } i \in C \tag{19a}$$

$$e_i^S = e_{ii}^S + \sum_{\substack{j=1,\\ j\neq i}}^{n} e_{ij}^S, \qquad \text{for } i \notin C. \tag{19b}$$

in accordance with substituting the equilibrium prices

$$\left(e_i^S\right)^* = e_i^S(p_1^*, \dots, p_n^*). \tag{20}$$

With respect to the world market for good X the equilibrium is obtained by simply replacing the equilibrium quantities of the decomposed dirty good from (10a) and (10b) into the production possibility frontier:

$$\left(x_i^S\right)^* = T\left(\left(e_i^S\right)^*\right), \qquad i = 1, \dots, n. \tag{21}$$

Combining the budget constraint (11) with the income functions (12a) and (12b) respectively[4], we get:

$$x_i^D = x_i^S + p_i e_{ii}^S + \sum_{\substack{j=1,\\ j\neq i}}^{n} p_j e_{ij}^S - p_i e_i^D + \left(\sum_{\substack{j=1,\\ j\in C,\\ j\neq i}}^{m} t_j e_{ij}^S - t_i \sum_{\substack{j=1,\\ j\in C,\\ j\neq i}}^{m} e_{ji}^S \right), \qquad \text{for } i \in C \tag{22a}$$

$$x_i^D = x_i^S + p_i e_{ii}^S + \sum_{\substack{j=1,\\ j\neq i}}^{n} p_j e_{ij}^S - p_i e_i^D, \qquad \text{for } i \notin C. \tag{22b}$$

Substituting (14), (21), and the equilibrium prices $p_1^*, \dots, p_n^*$ from (22a) and (22b) yields $\left(x_i^D\right)^*$ for both the coalition and the fringe countries:

$$\left(x_i^D\right)^* = x_i^D(e_1, \dots, e_n, t_1, \dots, t_n). \tag{23}$$

---

[4] By doing so, it can also be shown that Walras‘ Law holds for both the coalition and the fringe countries since, in (22a), the difference in parentheses which indicates the net tariff payment from coalition country i to the other coalition members $j \in C$, $j \neq i$ will be equal to zero. Hence, if all local fuel markets are in equilibrium, the world market for X must be in equilibrium as well.

The welfare function W_i of country i can now be derived by substituting (23) and the functional specifications from (6) and (8), in compliance with condition (14), into the utility function (5):

$$W_i(e_1, \dots, e_n, t_1, \dots, t_n) = V_i(e_i) + x_i^D(e_1, \dots, e_n, t_1, \dots, t_n) - D\left(\sum_{j=1}^{n} e_j\right) \tag{24}$$

which is, of course, different between coalition and fringe countries due to the variant term $x_i^D(.)$.

The benchmark for the subsequent analysis is the non-cooperative Nash equilibrium, i.e. the situation in which every country i chooses unilaterally a policy scheme (e_i, t_i) that maximizes the country's individual welfare function, taking as given the other countries' policy choices. We refer to this situation as the business-as-usual (BAU) scenario. The optimal BAU emission caps $(e_i)_{BAU}$ and tariff rates $(t_i)_{BAU}$ are obtained from the first-order conditions of the welfare maximization problem of country i:

$$\frac{\partial W_i}{\partial e_i} = \frac{\partial V_i}{\partial e_i} + \frac{\partial x_i^D}{\partial e_i} - \frac{\partial D}{\partial e_i} = 0 \tag{25}$$

$$\frac{\partial W_i}{\partial t_i} = \frac{\partial x_i^D}{\partial t_i} = 0. \tag{26}$$

Since countries face the same endowments and production technology, the BAU scenario is computed in a symmetric Nash fashion, yielding for $n = 10$

$$(e_i)_{BAU} = \frac{a(2\alpha^* + 9\alpha_H)}{9(b + 20\delta)\alpha_H + 2\alpha^*(b + 20\delta + 2\alpha_H)} \tag{27}$$

$$(t_i)_{BAU} = \frac{2a\alpha^*\alpha_H}{9(b + 20\delta)\alpha_H + 2\alpha^*(b + 20\delta + 2\alpha_H)} \tag{28}$$

As a consequence of symmetry, countries choose the same emission caps and tariffs in the absence of a coalition which leads to identical production and consumption of the commodities.

3 The Stackelberg Game

The set-up of the Stackelberg game basically is composed of two stages. On the first stage countries are involved in a strategic policy game and on the second stage agents maximize their rents from the production and consumption of commodities. Agents on the second stage behave perfectly competitive by taking into account any given policy measures. Governments on the first stage behave strategically as they aim to respond to the policy measures applied by all the other countries in an optimal way. Moreover, they try to anticipate the impact of their respective policies on the decisions of agents in the market who are confronted with these policies. In particular, we assume that members of a free trade area can coordinate their policies in terms of maximizing joint welfare. However, this does not necessarily mean that policies have to be harmonized. In any case coalition countries can employ the first-mover advantage of a Stackelberg leader.

With regard to the second stage, the stage of the various global and local markets, we already have determined the respective equilibrium quantities and equilibrium prices in the last section. By making use of these results we are able to turn to the policy game. First, we have to compute the welfare function of a coalition country as well as the one of a non-coalition country. Welfare generally is given by the utility of the consumer net of environmental damages since all other kinds of income generated in the economy are reflected in the consumer's budget.

Formally, the welfare functions for coalition and fringe countries, respectively is given by

$$W_i(e_1,\dots,e_n,t_1,\dots,t_n) = ae_i^D(e_i,t_i) - \frac{b}{2}\left(e_i^D(e_i,t_i)\right)^2 + x_i^D(e_1,\dots,e_n,t_1,\dots,t_n) - D\left(\sum_{j=1}^n e_j\right). \quad (29)$$

which can be specified as $W_{i\in C}$ for the coalition countries and as $W_{i\notin C}$ for the fringe countries by replacing the equilibrium quantities from (14) for e_i^D, $i = 1,\dots,n$ as well as x_i^D from (23).

For the solution of the Stackelberg Game, as a first step, the welfare of any fringe country has to be maximized with respect its cap and its tariff rate, taking the policies of all other countries as given. In this respect, fringe countries are viewed as behaving as non-cooperative Nash players facing the optimization problem

$$max_{e_i,t_i} W_{i\notin C}(e_1,\dots,e_n,t_1,\dots,t_n). \quad (30)$$

Differentiating the welfare function above with respect to the policies e_i, t_i gives the first-order conditions:

$$\frac{\partial W_{i\notin C}}{\partial e_i} = 0, \qquad \frac{\partial W_{i\notin C}}{\partial t_i} = 0. \quad (31)$$

Solving the first-order conditions for e_i, t_i yields the response function of a fringe country $i \notin C$ with respect to the policies of other fringe as well as coalition countries

$$\mathcal{R}_{e_i} = e_i(e_{-i},t_1,\dots,t_n) \quad (32)$$

$$\mathcal{R}_{t_i} = t_i(e_1,\dots,e_n,t_{-i}) \quad (33)$$

where $e_{-i} = (e_1,\dots,e_{i-1},e_{i+1},\dots,e_n)$ and $t_{-i} = (t_1,\dots,t_{i-1},t_{i+1},\dots,t_n)$.

For maximizing the welfare of coalition countries a different approach has to be taken. First of all, coalition countries can take advantage of anticipating how fringe countries will react to their strategies. That is how the fringe reaction functions enter the welfare of coalition countries. Secondly, as members of a coalition, in a sense, they take account of the impact of their policies on the group welfare of the coalition by internalizing any externalities they may impose, leading to the following optimization problem:

$$max_{e_1,\dots,e_m,t_1,\dots,t_m} \sum_{i=1}^{m} W_c(e_1, \dots, e_n, t_1, \dots, t_n). \qquad (34)$$

Replacing the response functions of the fringe countries for the fringe policies in the welfare functions of the coalition members and aggregating yields the joint welfare of the coalition

$$W_C = \sum_{i=1}^{m} W_{i\in C}\, (e_1, \dots, e_m, \mathcal{R}_{e_{m+1}}\big(e_{-(m+1)}, t_1, \dots, t_n\big), \dots, \mathcal{R}_{e_n}(e_{-n}, t_1, \dots, t_n),$$

$$t_1, \dots, t_m, \mathcal{R}_{t_{m+1}}(t_{-(m+1)}, e_1, \dots, e_n), \dots, \mathcal{R}_{t_n}(t_{-n}, e_1, \dots, e_n)) \qquad (35)$$

4 Self-enforcing IEAs

So far, we have dealt with an arbitrary, exogenously given number of coalition countries within the Stackelberg leader-follower framework. In doing so, considerations on the endogenous formation of a coalition and its stability have been omitted from the analysis. Hence, we need to examine which one of the potential coalition sizes $m \in [0, n]$ assures for a stable cooperation among member countries, or, put differently, which one constitutes a self-enforcing IEA. In the non-cooperative IEA literature, the notion of stability has proven to be a canonical requirement for environmental treaties to ensure the long lasting existence of climate coalitions of a particular size.[5] *Carraro and Siniscalco* (1993) put forth the profitability as a minimum requirement for coalition formation, although this does not prevent countries from free-riding. A coalition of size m is defined to be profitable if it brings about a welfare gain for its members compared to their BAU situation:

$$W_{i\in C}(m) \geq (W_i)_{BAU} \qquad (36)$$

In our framework, the stability requirement is met if a coalition of a certain size is found to be both internally and externally stable. In this respect, a coalition country i is defined to be internally stable if it does not have an incentive to leave the coalition C of size m, that is, if the following condition holds:

$$W_{i\in C}(m) \geq W_{i\notin C}(m-1) \qquad (37)$$

---

[5] The stability concept was originally elaborated by *d'Aspremont et al.* (1983) for the analysis of cartel formation in an oligopoly and later adapted to the IEA context (*Carraro and Siniscalco*, 1993, *Barrett*, 1994). The cooperative IEA literature introduced another notion of stability called the concept of the (gamma) core (*Finus*, 2003) which does not focus on individual player's strategies and payoffs along any coalition size but only on the countries' payoffs in the grand coalition compared to defection strategies such as joining sub-coalitions or unilateral free-riding (*Bréchet, Gerard and Tulkens*, 2011).

A coalition C of size m is defined to be externally stable if no fringe country i has an incentive to join the coalition, or, put formally, if

$$W_{i\notin C}(m) \geq W_{i\in C}(m+1) \tag{38}$$

Based on these considerations, all coalitions of integer size $m \in [0, n]$, that satisfy the equations $W_{i\in C}(m) - W_{i\notin C}(m-1) \geq 0$ and $W_{i\notin C}(m) - W_{i\in C}(m+1) \geq 0$ simultaneously, are found to be both internally and externally stable. As an illustration please consider the findings depicted in **Figure 1** and **Figure 2** which give the results of the simulation we have run for our model. The details of the parameterization as well as some additional results are presented in the next section. We find the stable coalition at size $m^* = 6$.

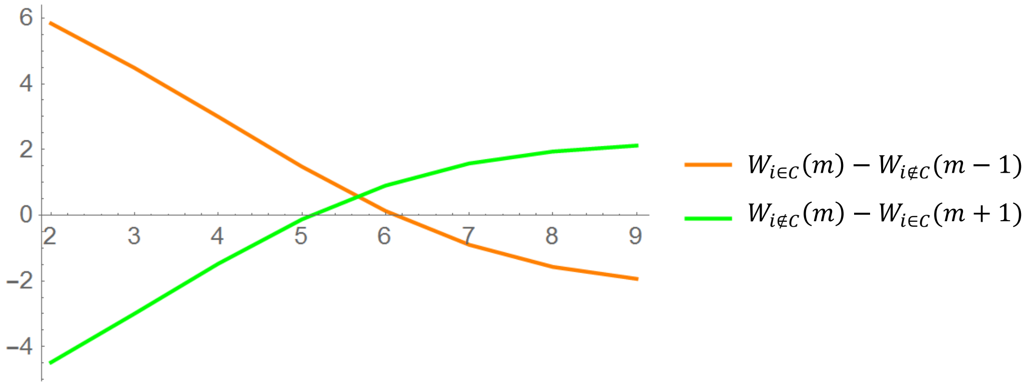


Figure 1: Internal and external coalition stability of the Stackelberg equilibrium with PFTA

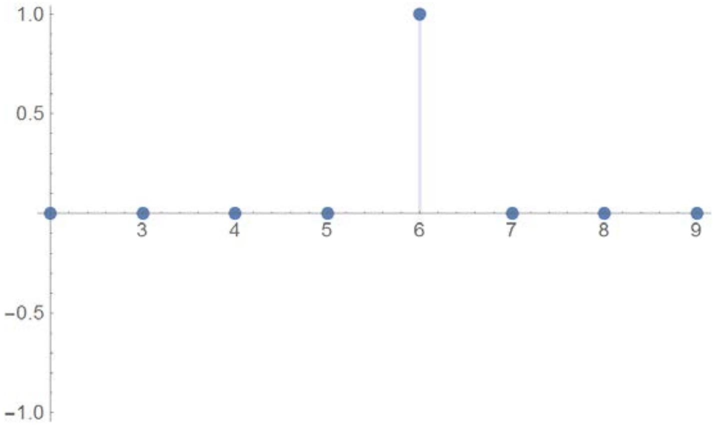


Figure 2: Stable coalition sizes of the Stackelberg equilibrium with PFTA

We should keep that in mind when the simulation results will be interpreted in the following section.

5 Results of the Simulation

Given the complexity of the model due to the fact that the market equilibria, welfare and response functions crucially depend on the endogenous coalition size, an analytical solution may be hard to compute, if not impossible. Nevertheless, in order to state some propositions on the role of the free trade arrangement for the formation of climate coalitions, we refer to numerical simulations exemplified by a run with the parameter values: $a = 100$, $b = 20$, $\bar{x} = 20$, $\alpha_H = 2000$, $\alpha^* = 2200$, $\delta = 10$ and $n = 10$. We consider a variation in coalition sizes in the range of $m \in [1,10]$ to examine its impact on emissions, damages and welfare. In each case the results are compared with the results of the BAU scenario. However, we should keep in mind that only a coalition of endogenous size $m^* = 6$ turns out being stable. That is why in the following the results for this particular coalition receive our special attention.

First of all, let us have a look at the total reduction of emissions achieved by the climate coalition viewed against the BAU scenario. As can be seen in **Figure 3**, coalitions of size $m \geq 3$ reduce total emissions significantly, the more the bigger is the coalition. For the stable coalition the reduction amounts to roughly 40 percent.

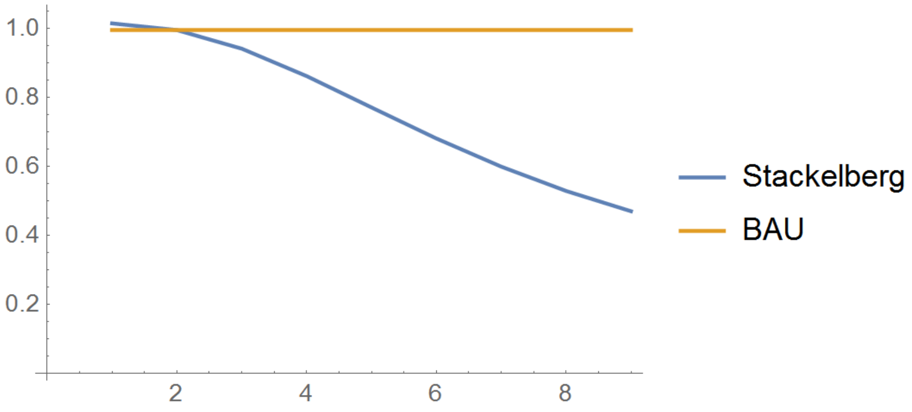


Figure 3: Total emissions in the Stackelberg equilibrium with PFTA

Figure 4 shows the emission caps set by each individual fringe and coalition country dependent on the coalition size.

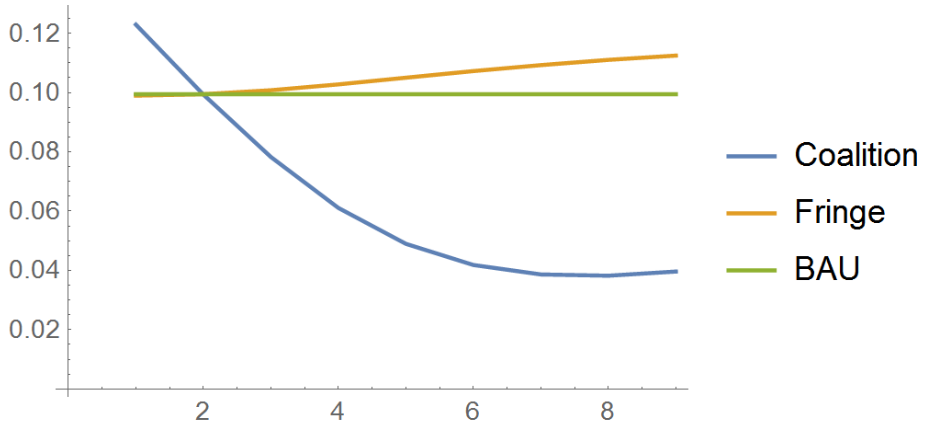


Figure 4: Individual caps in the Stackelberg equilibrium with PFTA

Apparently, as the size of the coalition is increasing, member countries reduce their individual emissions considerably, by about two thirds in case of the stable coalition. Even more surprising, fringe countries hardly increase their emissions, but only slightly, which explains why the coalition is fairly effective in combatting global warming and in preventing fringe countries from free-riding. The reason for that is to be found in a shift to the consumption of the clean good away from the consumption of the dirty good. In this sense, more imports of the dirty goods from fringe into coalition countries to replace for domestic production obviously have been successfully discouraged by means of the tariff. Again, this is a clear indication of the effectiveness of the free trade agreement in the mitigation of global warming.

The reduction in climate damages brought about by the emission reduction can be seen in **Figure 5**:

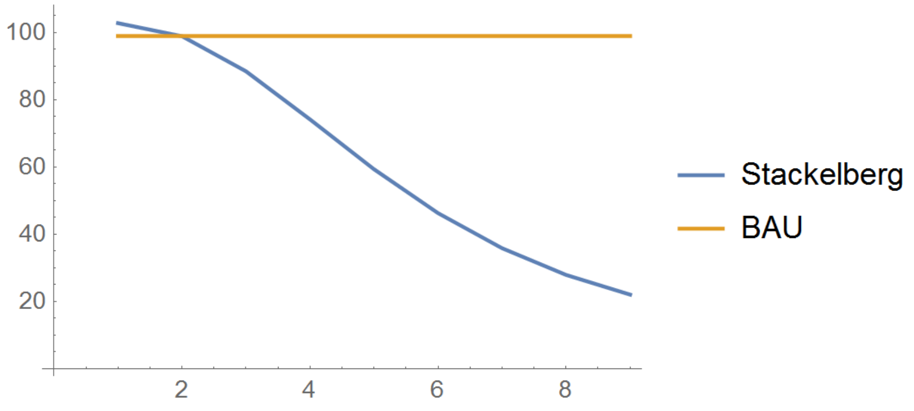


Figure 5: Total climate damages in the Stackelberg equilibrium with PFTA

In the stable case damages are reduced by about 60 percent compared to the BAU scenario for the quadratic damage function assumed.

The previous findings raise the issue of the induced change in welfare in the view of individual countries as well as from a global perspective. Firstly, let us have a look on global welfare as shown in **Figure 6**.

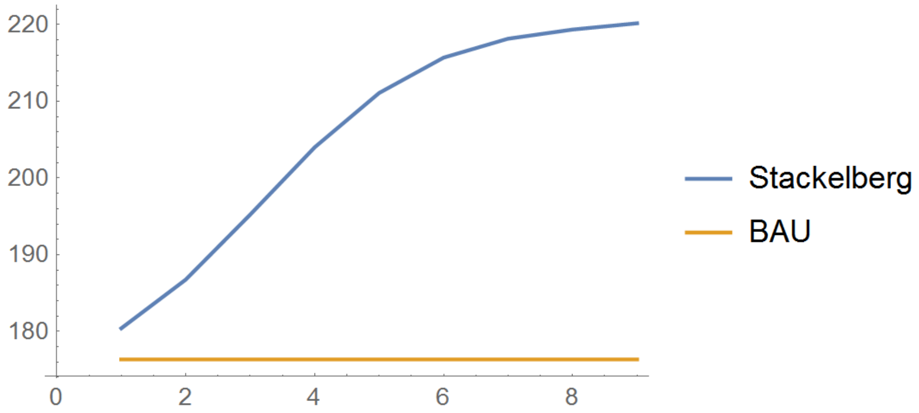


Figure 6: Total welfare in the Stackelberg equilibrium with PFTA

As can be seen, the free trade area as an incentive device for the formation of climate coalitions can lead to a considerable increase in world welfare relative to the BAU. In the stable case global welfare goes up by about one third. One might argue the reason for that may be primarily rooted in the trade liberalization prevailing in a free trade area in the usual welfare enhancing way. However, this is not the only reason as **Figure 5** indicates. Even more important is the reduction in global damages to the advantage of all countries provided by linking the climate change issue to the free trade arrangement.

Figure 7 shows how the welfares of the individual countries evolve dependent on the coalition size:

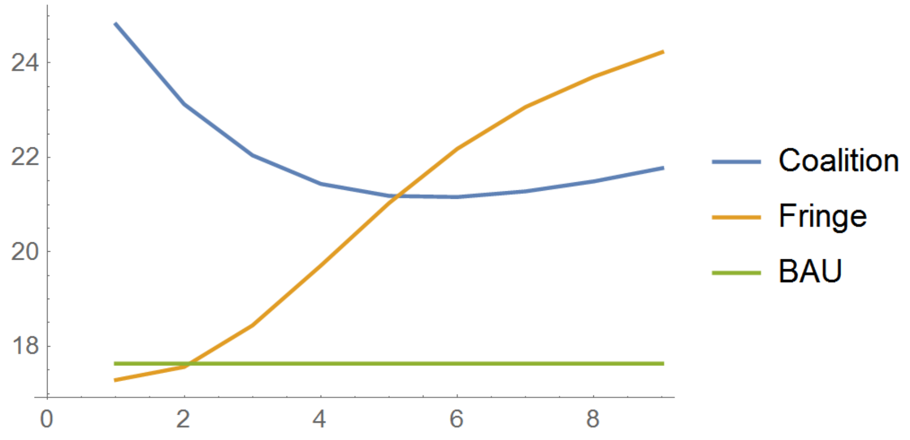


Figure 7: Individual welfare in the Stackelberg equilibrium with PFTA

Clearly, fringe countries would lose if they would join the coalition. This is the case primarily because they do not restrict the consumption of the dirty good as much as the coalition countries as **Figure 8** below is illustrating. There, consumption utility (i.e. welfare net of damages) is shown:

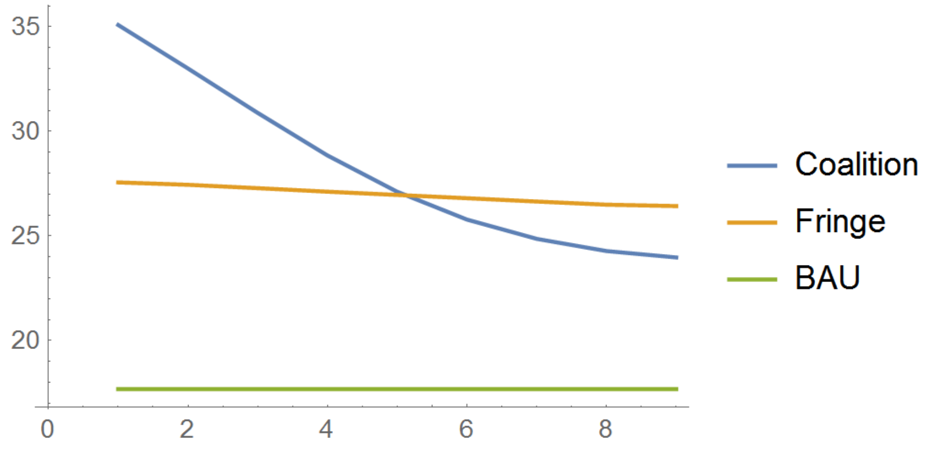
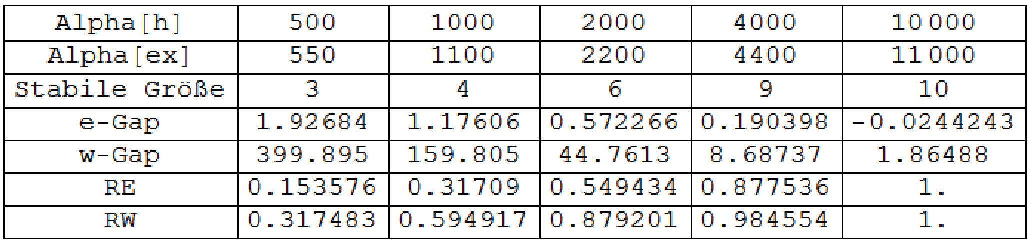


Figure 8: Consumption utility in the Stackelberg equilibrium with PFTA

However, countries as members of the stable coalition (at $m^* = 6$) gain tremendously compared to the option of free-riding.

A sensitivity analysis is carried out in the following, for variations in the parameters α_H and α^* as shown in **Table 1** below.



Alpha[h]	500	1000	2000	4000	10 000
Alpha[ex]	550	1100	2200	4400	11 000
Stabile Größe	3	4	6	9	10
e-Gap	1.92684	1.17606	0.572266	0.190398	-0.0244243
w-Gap	399.895	159.805	44.7613	8.68737	1.86488
RE	0.153576	0.31709	0.549434	0.877536	1.
RW	0.317483	0.594917	0.879201	0.984554	1.

Table 1: Variations of $\boldsymbol{\alpha_H}$ and $\boldsymbol{\alpha}^*$

The comparison of the different parameterizations is done using some effectiveness and efficiency measures proposed in the literature. The e-gap in **Table 1** measures the difference in global emissions between the BAU scenario and the social planer scenario. The latter is equivalent to the computation of the emissions in the Stackelberg-game with exogenous coalition size $n = 10$, that is, in the absence of any fringe countries. Hence, the emission gap is defined by:

$$e - gap = ne_{BAU} - ne_{SP} \tag{39}$$

The $w - gap$ measures the welfare gap between the BAU scenario and the social planer scenario:

$$w - gap = nw_{BAU} - nw_{SP} \tag{40}$$

The ratio RE measures the relative efficiency of emissions reductions of the climate coalition as it is defined as the emissions gap between BAU and the climate coalition in relative to the $e\text{-}gap$:

$$RE = \frac{ne_{BAU} - (me_C + (n-m)e_F)}{ne_{BAU} - ne_{SP}} \tag{41}$$

The ratio RW measures the relative welfare efficiency of the climate coalition in terms of the welfare gap between the coalition and BAU relative to the $w - gap$, i.e.

$$RW = \frac{(mw_C + (n-m)w_F) - nw_{BAU}}{nw_{SP} - nw_{BAU}} \tag{42}$$

As shown by **Table 1**, the stable coalition size increases as the parameters α_H and α^* increase. Increasing parameters α_H and α^* indicate an increase in the opportunity cost of the dirty good. Moreover, the advantage of the social planer scenario over the BAU scenario is diminishing since the dirty good is already consumed less in the BAU scenario due to its cost intensive production. This can be seen by the decreasing e-gaps and w-gaps along with α_H and α^*. At the same time, the efficiency of the climate coalition is increasing as measured by

RW and RE. The more expensive is the production of the dirty good, the more efficient is a free-trade based climate coalition.

6 Concluding Remarks

This paper addresses the role of trade liberalization on the endogenous formation of a self-enforcing climate coalition. We propose linking climate negotiations to negotiations on a PFTA while strategic trade and environmental policies vis-à-vis non-signatories are carried out individually. By doing so, we want to examine how the benefits resulting from the preferential trade liberalization affects the size, effectiveness, and stability of the climate coalition.

The model applied is an extension and modification of the Stackelberg leader-follower framework by *Eichner and Pethig* (2013a, 2013b) in which countries have two policy instruments at their disposal to strategically influence greenhouse gas emissions. On the one hand, they can discourage greenhouse gas emissions by means of an import tariff on dirty goods. They can, on the other hand, set an emissions cap affecting a national permit market. In order to identify the trade pattern within the free trade area and strictly distinct it from the trade patterns existing outside the area, we introduce a novel differentiation of firms' supplies with respect to the target markets. This implies a modification of the equilibrium concept for local markets.

The main focus is on the exploitation of terms of trade effects provided to members of the climate and free trade coalition that lets the PFTA become an incentive device for the formation of climate coalitions. The parametrical simulation shows evidence that the welfare gains provided by linking the IEA to the PFTA improve not only the effectiveness of the climate coalition in terms of emission reductions but also the stability by offsetting free-riding incentives, entailing a stable coalition size of $m^* = 6$. Global emissions as well as climate change damages are found significantly reduced in the numerical simulation of the model compared to the BAU scenario, while global welfare is found growing. Consequently, issue linkage with trade liberalization has the potential to promote and sustain broad international cooperation on climate change.

References

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

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

Barrett, S. (1997a): Heterogeneous international environmental agreements, in: Carraro, C. (ed.), *International Environmental Negotiations*, Edward Elgar Cheltenham, 9-25.

Barrett, S. (1997b): The strategy of trade sanctions in international environmental agreements, *Resource and Energy Economics* 19, 345-361.

Bosetti, V., Carraro, C., De Cian, E., Masetti, E. and Tavoni, M. (2013): Incentives and stability of international climate coalitions: An integrated assessment, *Energy Policy* 55, 44-56.

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

Bréchet, T., Gerard, F. and Tulkens, H. (2011): Efficiency *vs.* Stability in Climate Coalitions: A Conceptual and Computational Appraisal, *The Energy Journal* 32, 49-75.

Bucher, R. and Schenker, O. (2011): On Interactions of Optimal Climate Policy and International Trade: An Assessment of Border Carbon Measures, NCCR Research Paper No. 2011/11.

Carraro, C. (1999): The structure of international environmental agreements, in: Carraro, C. (ed.), *International Environmental Agreements on Climate Change*, Kluwer Dordrecht, 9-25.

Carraro, C. and Siniscalco, D. (1991): Strategies for the international protection of the environment, CEPR Discussion Paper No. 568.

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

Carraro, C. and Siniscalco, D. (1997): R&D cooperation and the stability of international environmental agreements, in: Carraro, C. (ed.), *International Environmental Negotiations*, Edward Elgar Cheltenham, 71-96.

Eichner, T. and Pethig, R. (2013a): Self-enforcing environmental agreements an international trade, *Journal of Public Economics* 102, 37-50.

Eichner, T. and Pethig, R. (2013b): Trade tariffs and self-enforcing environmental agreements, CESifo Working Paper No. 4464.

Eyckmans, J. and Finus, M. (2007): Measures to enhance the success of global climate treaties, *International Environmental Agreements* 7, 73-97.

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

Finus, M. and Rundshagen, B. (2000): Strategic Links between Environmental and Trade Policies if Plant Location is Endogenous, University of Hagen Working Paper No. 283.

Fischer, C. and Fox, A.K. (2012): Comparing policies to combat emissions leakage: Border carbon adjustments versus rebates, *Journal of Environmental Economics and Management* 64, 199-216.

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

Kemfert. C. (2004): Climate coalitions and international trade: assessment of cooperation incentives by issue linkage, *Energy Policy* 32, 455-465.

Leal-Arcas, R. (2013): Climate Change Mitigation from the Bottom Up: Using Preferential Trade Agreements to Promote Climate Change Mitigation, *Carbon and Climate Law Review* 7, 34-42.