Graz Economics Papers – GEP

GEP 2020–04

Measures to Enhance the Effectiveness of International Climate Agreements: The Case of Border Carbon Adjustments

Alaa Al Khourdajie and Michael Finus

February 2020

Department of Economics
Department of Public Economics
University of Graz

Measures to Enhance the Effectiveness of International Climate Agreements: The Case of Border Carbon Adjustments

Alaa Al Khourdajie* and Michael Finus†‡

ABSTRACT

Actions on climate change which are not supported by all countries are not very effective. However, full participation in a global climate treaty with meaningful emission reductions is difficult to achieve. The non-excludability of the public good mitigation provides an incentive to abstain from global action. Moreover, carbon leakage renders it unattractive to join a treaty without full participation. We study whether and under which conditions border carbon adjustments (BCAs) can mitigate free-riding and reduce carbon leakage in a simple strategic trade model. We show that BCAs can lead to large stable climate agreements, including full participation, associated with large global welfare gains if treaties do not restrict membership (open membership), as this is typical for environmental agreements. We caution against restricting accession to treaties (exclusive membership), as this is typical for trade agreements, which may serve individual but not global interests.

Keywords: *Self-enforcing international climate agreements, International trade, Border carbon adjustments*

JEL-Classification: C71, D62, F18, H23, H41, Q54.

*Centre for Environmental Policy, Imperial College London, London, SW7 2AZ, UK

†Department of Economics, University of Graz, Austria and University of Bath, UK.

‡Corresponding author at: Department of Economics, University of Graz, Austria and University of Bath, UK.

E-mail addresses: A.Alkhourdajie@imperial.ac.uk (A.AlKhourdajie), Michael.Finus@uni-graz.at (M.Finus).

1. Introduction

Climate change actions which do not receive the support by all countries are not very effective. However, full participation in a global climate treaty with meaningful emission reductions is difficult to achieve. Due to the non-excludability of the public good mitigation, non-signatories benefit from emission reductions of signatories to a climate agreement without incurring abatement costs. This provides a free-rider incentive. Moreover, there is a relocation of production from "clean" signatory to "dirty" non-signatory countries. This carbon leakage effect and loss of competitiveness render a climate treaty less effective and discourage participation. In order to reduce free-rider incentives and carbon leakage, trade measures such as border carbon adjustments (BCAs) have become increasingly popular in the policy debate (e.g., Financial Times, Feb. 12, 2017; Zenghelis and Stern, 2009) but also in academic circles (e.g., Böhringer et al., 2017b; Fischer and Fox, 2012; Keen and Kotsogiannis, 2014; Nordhaus 2015).

The idea of BCAs[1] is simple. Suppose members to a climate treaty pursue a more ambitious climate policy than non-members. Environmental policies impose costs on firms, also called carbon prices.[2] If countries implement an emission tax t for instance, different taxes translate into different carbon prices: $t_S > t_{NS}$ with subscript S denoting signatories and NS non-signatories, respectively. One measure of BCAs is that signatories impose a tariff τ on imports such that foreign firms face not only t_{NS} but $t_{NS}+\tau$ whenever they export to signatory markets. If $t_{NS}+\tau = t_S$, this is a fully adjusted import tariff and if $t_{NS} + \tau < t_S$, this would be a partially adjusted import tariff. An adjustment $t_{NS} + \tau > t_S$ would violate the principle of equal treatment under the regulations of WTO. With a fully adjusted import tariff, firms play on equal terms in signatory markets. However, this is not the case in non-signatory markets. Hence, signatory governments could grant their firms an export rebate κ on their exports to non-signatory markets, such that their firms pay only $t_S - \kappa$ on exports. A full rebate implies $t_S - \kappa = 0$ and a fully adjusted rebate implies $t_S - \kappa = t_{NS}$ (whereas $t_S - \kappa < t_{NS}$ would violate WTO rules).[3] Most authors talk about full BCAs if import tariffs are combined with exports rebates.

Both tariffs and rebates aim at creating an equal playing field, reducing the disadvantage of firms located in signatory countries due to stricter environmental

[1] We use the term border carbon adjustements (BCAs) to stress the environmental justification of border tax adjustments (BTAs) in the context of this paper as BTAs are also discussed for other reasons, such as different labor standards and corporate taxes across countries.

[2] To be precise, the carbon price is equal to the marginal abatement cost of the last unit of emission reduction, which corresponds to the tax per unit of emission or the price of tradable emission permits, abstracting from any other distortion.

[3] Note that export rebates can be paid directly or indirectly. An example of indirect payment is the exemption of emission-intensive and trade exposed industries from strict environmental regulation, e.g., the steel and aluminum industry, within the European Union Emission Trading System (EU-ETS). Those industries received a large number of freely allocated tradable permits (Böhringer et al., 2017b).

regulation. Both instruments reduce the leakage effect, even though without rebates the emission reduction effect is larger. Export rebates are also more difficult to justify. The argument that emissions are harmful and therefore should be taxed, should hold true irrespective of whether firms produce for their domestic market or for foreign markets.

The academic discussion about BCAs has broadly focused on three issues. The first issue relates to the detailed design and practicability of trade measures for environmental reasons and to what extent they are compatible with WTO rules (e.g., Fischer and Fox, 2012; Mehling et al., 2019). It appears that the overarching conclusion is that, despite many practical obstacles regarding implementation, BCAs generally do not violate WTO rules.

The second issue relates to the economic justification of trade measures for environmental reasons.[4] Despite the fact that conventional trade theory argues against trade barriers (i.e., different environmental standards simply reflect different environmental preferences and/or comparative advantages related to the relative abundance of the environment and natural resources of countries), in the context of global pollutants, the correction of distortions through market interventions can be justified. The internalization through first- and second-best instruments has been analyzed for instance by Copeland and Taylor (1995), Markusen (1975), Hoel (1996) and more recently by Keen and Kotsogiannis (2014), Vlassis (2013) and Tsakiris (2014) for example. It is also for this reason that Stiglitz (2006) argues that the absence of carbon prices de facto constitutes a subsidy for dirty production. Hence, BCAs are correcting a market imperfection by internalizing the social cost of carbon. However, it is also clear that if BCAs do not enforce uniformly higher environmental standards at the global scale, but only among a subgroup of countries, the welfare gains from a tougher environmental policy may be small if not negative as BCAs may seriously harm outsiders. In other words, in an ideal world, BCAs are a threat with which to enforce full cooperation (or something close to this), but if they are successful then they are not implemented.

The third issue relates to the effectiveness of trade measures for environmental reasons, drawing on two strands of literature. The first strand, which comprises the bulk of papers, conducts numerical simulations with empirically calibrated computable general equilibrium (CGE) models in order to estimate the reduction of leakage effects through various forms of BCAs at the aggregate level and in particular sectors (Böhringer et al., 2014, 2015, 2017a,b; Caron 2012; Fischer and Fox, 2012; see also Branger and Quirion, 2014 for a meta-analysis of 25 studies). Most of these studies confirm the conclusion that export rebates can reduce carbon leakage but only import tariffs lead to a noticeable overall reduction of emissions as they also curtail consumption. Due to the complexity of these multi-country, multi-sector models, the environmental target of signatory countries is set exogenously and the environmental targets of non-signatories is fixed at the level without BCAs. That is, the strategic interaction between signatories and non-signatories in a game-theoretic

[4] For an excellent discussion of this issue, see for instance Helm et al. (2012).

sense and the resulting endogenous policy levels are not explicitly captured. Moreover, in contrast to the original idea of BCAs, namely to enhance the effectivness of climate treaties, the formation and stability of climate treaties is not tested.[5]

The second strand, which represents a much smaller part of the literature on the third issue, is theoretical in nature. It focuses on the endogenous choice of policy levels among countries in a strategic trade model of imperfect competition. Surprisingly, this literature has also mainly ignored the issue of treaty formation by restricting attention to two countries (e.g., Anoulis, 2014; Baksi and Chaudhuri, 2017; Eyland and Zaccour, 2012 and 2014). Typically, these models consider a home and a foreign country with a different perception of environmental damages such that it is not profitable for the foreign country to embark on a fully cooperative solution. BCAs are then considered as a threat with which to enforce cooperation. As far as we are aware, only Baksi and Chaudhuri (2014) consider BCAs in n-country model of coalition formation. However, they only analyze the stability of the grand coalition and ignore the entire process of treaty formation. Also, Barrett (1997) considers coalition formation in a strategic trade model. But his model, as he himself admits, is not very general. He sets non-signatories' abatement level to zero, considers only the option of a complete trade ban and supports stability through a minimum participation clause, which, according to the body of literature, stabilizes large coalitions by itself.

Regarding the third issue, our paper is in the tradition of the game-theoretic literature. We model the entire process of treaty formation including stability, but admittedly the model is highly stylized. Our model directly benefits from three sets of papers. First, the literature on strategic trade models, which extended the model by Brander and Spencer (1985) by including environmental damages and consumer surplus in governments' welfare function (e.g., Burguet and Sempere, 2003; Baksi and Chaudhuri, 2009; Barrett, 1994b; Conrad, 1993; Kennedy, 1994; Ulph 1996), but this literature has mainly restricted attention to only two countries and/or ignored the issue of agreement formation. Second, the literature on international environmental agreements, which focuses on treaty formation and emission reduction, but typically ignores trade.[6] Third, the literature on trade agreements, but without considering environmental issues, from which it emerges that membership rules may determine the success of treaty formation (see in particular Yi, 1996, 2000; see also Loke and

[5]An exception is Irfanouglu et al. (2015). However, their analysis is restricted to three countries. Although conceptually interesting, Helm et al. (2012) remain at a rather stylized level in their analysis. Another exception is Weitzel et al. (2012). However, they only test the stability of the grand coalition (i.e., a coalition including all countries) and the group of Annex 1 countries (i.e., those 37 countries which have accepted emission ceilings under the Kyoto Protocol signed in 1997). Also Nordhaus (2015) analyzes coalition stability but policy levels are exogenous.

[6]See the survey of this literature by Finus and Caparrós (2015), including a collection of the most influential articles since the early papers by Barrett (1994a) and Carraro and Siniscalco (1993). Eichner and Pethig (2013, 2015) have recently introduced trade in the literature on internatonal environmental agreements, though their model is very different from our intra-industry trade model and they do not consider BCAs.

Winters, 2012; Helpman and Krugman, 1985).

This paper builds on Finus and Al Khourdajie (2018), who consider taxation in an intra-industry trade model with horizontal product differentiation, taste for variety (hereinafter abbreviated as TFV) by consumers and coalition formation, but without BCAs. This serves as a good benchmark to analyze the effect of BCAs in the form of import tariffs. We do not consider export rebates for two reasons. First, we show that even BCAs without export rebates support successful climate agreements. Second, export rebates are not in line with the idea that harmful production should be regulated irrespective of the market for which goods are produced. Hence, without rebates, a justification of BCAs within WTO-rules is much easier and a justification of possible retaliatory measures by non-signatories is much weaker. We briefly discuss alternative assumptions in the concluding section 6.

In this paper, we are interested in two main issues.

First, how do BCAs change the incentives of an n-country public good provision game with the possibility of forming treaties as modelled by the literature on international environmental agreements? We show that without BCAs stable agreements and the gains from cooperation are small, in line with the "paradox of cooperation", a term coined by Barrett (1994) or what Nordhaus (2015) called the "small coalition paradox". We demonstrate that BCAs are a "game changer". On the one hand, countries find it less attractive to stay outside due to the negative impacts of BCAs on non-signatories. On the other hand, countries find it more attractive to join an agreement, as the gains from cooperation are larger due to lower leakage effects.

Second, we are interested in how different membership rules affect the success of a climate treaty with BCAs. Hence, we compare the outcome of an agreement under open and exclusive membership. We show that large and therefore successful stable agreements emerge under open membership, as a type of "large coalition anti-paradox". In contrast, under exclusive membership, we demonstrate that signatories may find it attractive to restrict accession to their agreement for selfish motives before large or full participation is reached. Consequently, as BCAs punish non-participation, not surprisingly, agreements without large participation may be associated with large global welfare costs. Accordingly, in order to avoid this problem, climate treaties should allow for accession without restricting membership. This would be very much in line with most international environmental treaties, though different from most trade agreements, which typically condition accession of new members on an approval procedure.

Our paper proceeds as follows. In section 2, we present the three-stage coalition formation model in which countries first choose their membership, then they choose their policy levels and finally firms choose their output. We solve the game by backward induction and hence consider the output stage in section 3, the policy stage in section 4 and the membership stage, including an overall evaluation of stable agreements, in section 5. Section 6 concludes, qualifies our results and points to future research.

2. Model

2.1. Coalition Formation Game

Consider an intra-industry trade model with n *ex-ante* symmetric countries having a representative firm and consumer in each country. We denote the set of countries by N with n the number of countries, the cardinality of N. Firms produce a horizontally differentiated good, i.e., the same good but in different varieties where each firm produces one variety. The production of this good releases greenhouse gases which cause environmental damages. Firms compete in a Nash-Cournot fashion. Markets are segmented and each firm supplies its good to the domestic and all foreign markets. Because of the segmentation of markets, firms play a separate Cournot-game in each market. Transport costs are assumed away as usual.

We assume a three-stage coalition formation game:

Stage 1, Choice of Membership: all countries decide simultaneously and non-cooperatively whether to join coalition $S \subseteq N$ with m the cardinality of S, $1 \leq m \leq n$. Countries which do not join S act as singletons. A representative signatory will be denoted by i and a representative non-signatory by j.

Stage 2, Choice of Policy Level: all countries choose simultaneously their emission tax.

- Signatories choose their tax t_i (implemented uniformly in all signatory countries) by maximizing the joint welfare of coalition S: $\underset{t_i}{max} \sum_{i \in S} W_i$.
- Non-signatories choose their individual tax t_j by maximizing their individual welfare: $\underset{t_j}{max}\, W_j$.

Stage 3, Choice of Output: all firms choose simultaneously and non-cooperatively their segmented market outputs for each of the n markets by maximizing their producer surplus: $\underset{q_{1i},\cdots,q_{ni}}{max}\ PS_i$.

The game is solved by backward induction.

In the third stage, firms play a Nash equilibrium output game in each of the n segmented markets. Output q_{ki} refers to the output of firm i sold in market k. Firm i's output q_{ki} depends on the demand in market k, production costs and taxes. We consider two tax regimes.

The first tax regime is called the *No BCA-regime.* Each government imposes a tax on its representative firm. Firms in signatory countries face tax t_i and firms in non-signatory countries face tax t_j. As will become clear below, the emissions tax is equivalent to an output tax as the tax is imposed per unit of output and we assume a constant emission output ratio which we normalize to one without loss of generality.

The second tax regime is called the *BCA-regime.* In terms of the tax structure (though not in terms of equilibrium taxes), it is almost identical to the first regime, except for firms in non-signatory countries. No changes apply to quantities sold to their own market or any other non-signatory market; they face tax t_j. Also no changes apply to quantities sold to signatory markets if $t_j \geq t_i$; they face t_j. However, changes apply if $t_j < t_i$; then they need to pay additionally a mark-up $\phi(t_i - t_j)$, i.e., the border carbon adjustment. Hence, firms pay $t_j + \phi(t_i - t_j)$ per unit of output.

Following Eyland and Zaccour (2012), we call ϕ the BCA-adjustment parameter. In accordance with the anti-discrimination rules of the WTO, $\phi > 1$ is not feasible. Hence, we assume that $0 \leq \phi \leq 1$. Obviously, BCA- and No BCA-regime coincide for $\phi = 0$.

Taken together, the third stage delivers an equilibrium vector of outputs of firms located in signatory and non-signatory countries which is a function of all taxes.

In the second stage, inserting equilibrium outputs as a function of taxes from stage 3 in countries' welfare functions, we can solve for countries' equilibrium tax rates. Signatories, as a group, choose their taxes by acting as one single player whereas each non-signatory acts as a single player. Hence, the solution of the second stage can be interpreted as a Nash equilibrium in a tax game between coalition S as one player and $n - m$ singleton players. As we will see, equilibrium taxes depend on the size of coalition S, m, and whether border carbon adjustments are available to signatories. If we insert equilibrium tax rates (stage 2) into equilibrium outputs (stage 3) and those outputs into welfare functions, welfare of signatories and non-signatories can be expressed as a function of coalition size m, $W_{i \in S}(m)$ and $W_{j \notin S}(m)$, respectively.[7] This provides the input for the first stage.

In the first stage, we solve for a Nash equilibrium in membership strategies. We consider a simple cartel formation game (d'Aspremont et al., 1983), also called a simultaneous move open-membership single coalition game (Yi, 1997) in order to stress the institutional setting of this type of agreement. Each country simultaneously chooses whether to join coalition S or to remain a singleton. The treaty is of the open-membership type: every country can join coalition S if it wishes to do so. A coalition is called stable if no country has an incentive to change its announcement, given the announcements of all other countries. Following d'Aspremont et al. (1983), we can also say that no signatory i has an incentive to leave coalition S to become a non-signatory (internal stability) and no non-signatory j has an incentive to join coalition S to become a signatory (external stability):

- Internal Stability: $W_{i \in S}(m) - W_{i \notin S}(m-1) \geq 0$
- External Stability: $W_{j \notin S}(m) - W_{j \in S}(m+1) \geq 0$.

The size of a coalition which is internally and externally stable and hence stable is denoted by m^*. In the course of the later discussion, we will also consider alternative institutional settings, e.g., a sequential choice of membership and exclusive membership.

[7] Due to the symmetry assumption, all signatories (non-signatories) will have the same welfare, though welfare levels will differ between both groups. That is, all countries are ex-ante symmetrical, although they are ex-post asymmetrical, because countries which join S will choose a different tax level than countries which do not belong to S.

We note that if the grand coalition forms (i.e., a coalition of all players; $m = n$), this replicates the social optimum, and if no coalition forms (i.e., all players act as singletons; $m = 1$), this replicates a Nash equilibrium in a game without coalition formation. Moreover, if the grand coalition forms and hence there is no outsider left in the game, the taxes under the BCA- and No BCA-regime coincide, which are socially optimal. Also if all players act as singletons and hence no agreement has formed, no tariff can be implemented at the border and taxes under both regimes coincide. It is for this reason that in many of the subsequent propositions and results, which compare signatories and non-signatories, we assume $1 < m < n$.

2.2. Welfare Function

In this subsection, we have a closer look at the welfare function of governments and its different components.

If a country i becomes a signatory, $i \in S$, its welfare is given by:

$$W_i = CS_i + PS_i - D_i + TR_i + BCR_i \tag{1}$$

where CS_i represents country i's consumer surplus, PS_i country i's producer surplus, D_i is the pollution damage faced by country i, TR_i is country i's tax revenue from the tax imposed on its domestic firm's production, and BCR_i is country i's tariff revenue from the tariff imposed on imports from firms located in non-signatory countries. Of course, under the No BCA-regime $BCR_i = 0$.

If a country j remains a non-signatory, $j \in N \backslash S$, its welfare is given by:

$$W_j = CS_j + PS_j - D_j + TR_j \tag{2}$$

with the same welfare components as in Eq. (1) after the appropriate changes of subscripts. Note that the BCR-term is missing in a non-signatory government's welfare function as, per assumption, non-signatory governments cannot impose import tariffs. (See the discussion in section 6.)

The inverse demand function in country i for country k's variety is given by:

$$p_{ik} = a - (1-\gamma)q_{ik} - \gamma Q_{i.} \Longleftrightarrow p_{ik} = a - q_{ik} - \gamma \sum_{l \in N, l \neq k} q_{il} \tag{3}$$

where p_{ik} represents the price faced by consumers in country i consuming the variety supplied by a firm located in country k and q_{ik} is the corresponding quantity[8]; a is a positive demand parameter and $Q_{i.}$ is a short-hand notation which stands for $Q_{i.} = \sum_{k \in N} q_{ik}$, i.e., country i's total consumption of all varieties supplied by all firms k (located in signatory and non-signatory countries). Hence, the dot stands for the total sum over an index. The term $\sum_{l \in N, l \neq k} q_{il}$ is the sum of all consumed varieties produced by all firms except the firm located in country k.

[8]Throughout the paper the first subscript indicates the market in which the variety is consumed and the second subscript indicates the country in which it is produced.

Accordingly, the consumer surplus in country i is given by:

$$CS_i = aQ_{i.} - \frac{\gamma}{2}Q_{i.}^2 - \frac{1-\gamma}{2}\sum_{k\in N} q_{ik}^2 - \sum_{k\in N} q_{ik}p_{ik} \tag{4}$$

where the first three terms represent the utility derived from the consumption of the horizontally differentiated and traded good and the last term in Eq. (4) is the representative consumer i's expenditure. Thus, the consumer surplus is a function of all varieties $q_i = (q_{i1}, ..., q_{in})$ consumed by consumers in country i.

Consumers have a taste for variety (Dixit and Stiglitz, 1977). Hence, their utility depends not only on the total quantity consumed (term $aQ_{i.} - \frac{\gamma}{2}Q_{i.}^2$ in Eq. (4)) but also on the composition of quantities of the differentiated good (term $-\frac{1-\gamma}{2}\sum_{k\in N} q_{ik}^2$ in Eq. (4)) as for instance assumed in Yi (1996, 2000). The taste for variety (TFV) is captured by the parameter $\gamma \in [0,1]$. High values of γ imply a low taste for variety and low values of γ correspond to a high taste for variety. For $\gamma = 1$ varieties are perfect substitutes and for $\gamma = 0$ varieties cannot be substituted at all.[9] Introducing a taste for variety allows us to capture the observation that for many goods, varieties are not always perfect substitutes. The larger the value of γ, the higher will be the competition among firms and, hence, the stronger will be the strategic interaction among governments. For $\gamma = 0$ firms act like a monopolist for their variety in the market. It will turn out that the value of γ crucially affects the success of agreement formation.

For producers, allowing for the possibility of border carbon adjustments, we need to distinguish between firms located in signatory and non-signatory countries. The producer surplus of a firm located in a signatory country i is the sum of its profit obtained in each market:

$$PS_i = \sum_{k\in S} \pi_{ki} + \sum_{l\in N\setminus S} \pi_{li} = \sum_{k\in S} q_{ki}(p_{ki} - c - t_i) + \sum_{l\in N\setminus S} q_{li}(p_{li} - c - t_i) \tag{5}$$

where π_{ki} (π_{li}) represents firm i's profit in signatory k's (non-signatory l's) market from selling quantity q_{ki} (q_{li}) at price p_{ki} (p_{li}); c is a constant marginal cost parameter and t_i is the tax imposed by country i's government on its firm's production.

Also the producer surplus of a firm located in a non-signatory country j is the sum of its profit in each market:

$$PS_j = \sum_{k\in S} \pi_{kj} + \sum_{l\in N\setminus S} \pi_{lj} = \sum_{k\in S} q_{kj}(p_{kj} - c - t_j - \Omega) + \sum_{l\in N\setminus S} q_{lj}(p_{lj} - c - t_j) \tag{6}$$

$$with \;\; \Omega = \begin{cases} \phi(t_i - t_j) & if \;\; t_i > t_j \\ 0 & if \;\; t_i \leq t_j \end{cases}$$

[9] An extension could be the "ideal variety" approach where consumers have not only a general preference for the variety of a good but also a preference for a particular variety. One application is a preference for the domestically produced variety (Di Comite et al., 2014).

where π_{kj} (π_{lj}) represents firm j's profit in signatory k's (non-signatory l's) market from selling quantity q_{kj} (q_{lj}) at price p_{kj} (p_{lj}), and t_j is the tax imposed by country j's government on its firm's production. Clearly, under the No BCA-regime, $\phi = 0$, and, therefore, $\Omega = 0$. Hence, the structure of the producer surplus of a firm located in a signatory and non-signatory country is the same. Under the BCA-regime, this is also the case if $t_i \leq t_j$ because then $\Omega = 0$. However, this is different if $t_i > t_j$ because then $\Omega > 0$ provided there is some adjustment, i.e., $\phi > 0$. Then, firms located in non-signatory countries face tax $t_j + \phi(t_i - t_j)$ for all quantities which they export to signatory markets. For all other quantities, they face only their local tax t_j.

Damages in signatory and non-signatory countries are the same ($D_i = D_j = D_l$) and are given by:

$$D_l = \delta Q \tag{7}$$

where δ is a damage parameter, $Q = \sum_{i \in N} Q_{.i} = \sum_{i \in N} Q_{i.}$ is total production which is equal to total consumption worldwide. That is, we assume a constant emission output coefficient which we normalize to 1 without loss of generality. Hence, we assume that emissions constitute a pure public bad: damages depend on total emissions. As there is no abatement technology in our simple model, emission and output tax are the same and an output/emission tax is an efficient policy instrument to address externalities.

The tax revenue of signatory country i is given by:

$$TR_i = t_i \sum_{k \in N} q_{ki} \tag{8}$$

which is the tax rate it imposes on its firm multiplied by the total quantity produced by its firm for all markets. Similarly, the tax revenue of non-signatory country j is given by:

$$TR_j = t_j \sum_{k \in N} q_{kj} \ . \tag{9}$$

Finally, under the BCA-regime, the border carbon adjustment revenue of a signatory country obtained from the adjustment Ω on imports from firms located in non-signatory countries is given by:

$$BCR_i = \Omega \sum_{j \in N \setminus S} q_{ij} \tag{10}$$

with Ω defined in Eq. (6).

3. Third Stage

In this section, we analyze output of firms in the third stage. The details and all proofs are provided in Appendix 1. Each firm supplies all markets. Given the focus of this paper on agreement formation, with two groups of countries, i.e., signatories and non-signatories, markets can be divided into signatory and non-signatory markets. Generally, each firm's variety depends on own taxes and all foreign taxes. However, given that all governments have the same welfare function in our model, it will turn out in stage 2 below that all signatories $i \in S$ will choose the same tax t_i and all non-signatories $j \notin S$ the same tax t_j, though normally $t_i \neq t_j$, with a special interest in situations for which $t_i > t_j$ is true. Hence, it is convenient to view quantities already in stage 3 as being a function of signatories' taxes t_i and non-signatories' t_j taxes only. Henceforth, in order to render the comparison between the two regimes interesting, we assume $t_i > t_j$ under the BCA-regime, i.e., signatories choose a higher tax (or lower subsidy) than non-signatories (as otherwise the BCA- and No BCA-regime are identical).

We further assume for the adjustment parameter $\phi = 1$ such that non-signatories face de facto t_i on their exports to signatory countries under the BCA-regime if $t_i > t_j$ (because $t_j + \phi(t_i - t_j) = t_i$ if $\phi = 1$). The reason is that signatories would have no reason to choose any smaller value could they choose this parameter optimally and we recall that $\phi > 1$ would violate WTO rules. This equips signatories with the "maximum legally possible enforcement power".

Proposition 1 highlights the strategic effect of taxes on outputs under both regimes.

Proposition 1 - The Effect of Taxes on Equilibrium Production

Suppose a coalition S with m signatories has formed in the first stage, $1 < m < n$, and let $t_i > t_j$ under the BCA-regime.

1) *Let $\gamma > 0$.*

a) *Under the No BCA-regime, signatory firms' output for every market decreases in signatories' taxes and increases in non-signatories' taxes. Similarly, non-signatory firms' outputs decrease in non-signatories' taxes and increase in signatories' taxes.*

b) *Under the BCA-regime, this is also true for all outputs sold to non-signatory markets. However, all firms' outputs sold to signatory markets decrease in signatories' taxes but are independent of non-signatories' taxes.*

2) *Let $\gamma = 0$.*

a) *Under the No BCA-regime, signatory firms' outputs decrease in signatories' taxes but are independent of non-signatories' taxes. Similarly, non-signatory firms' outputs decrease in non-signatories' taxes but are independent of signatories' taxes.*

b) *Under the BCA-regime, this is also true for outputs sold to non-signatory markets; outputs sold to signatory markets decrease in signatories' taxes and are independent of non-signatories' taxes.*

As expected, under the No BCA-regime, quantities of a firm's variety are negatively affected by domestic taxes and positively affected by foreign taxes (Part 1a in Proposition 1).

Under the BCA-regime, the same strategic interaction applies to "unprotected" non-signatory markets. However, in the "protected" signatory markets, all firms face de facto the same tax t_i under the BCA-regime and therefore non-signatories' taxes t_j do not affect quantities supplied in these markets (Part 1b in Proposition 1).

Finally, for full TFV, i.e., $\gamma = 0$, a firm's output is not affected by foreign taxes under the No BCA-regime (Part 2a in Proposition 1). In this particular case, regarding its own variety, each firm acts like a monopolist in each market as consumers do not substitute between varieties. Hence, there is no competition among firms. That is, each firm's output is only negatively affected by the taxes of its own government. Under the BCA-regime, this only applies to non-signatory markets (Part 2b in Proposition 1). Due to the import tariff, signatories control their market and, hence, in signatory markets outputs only decrease in signatories' taxes, as this was already observed in Part 1b in Proposition 1 for $\gamma > 0$.

Taken together, compared to the No BCA-regime, under the BCA-regime, governments in signatory countries jointly enjoy more market power by protecting their domestic markets through import tariffs. That is, if signatory governments choose a higher tax than non-signatory governments under the BCA-regime, at least in signatory markets, all firms face de facto the same tax t_i and hence act on an equal playing field. Moreover, importantly, high (low) values of the TFV-parameter γ imply a high (low) degree of strategic interaction between signatories and non-signatories in this tax competition game.

The following proposition sheds further light on the difference between production and consumption patterns under both regimes. In order to focus the analysis, henceforth, we assume $t_i > t_j$, not only under the BCA-regime, but also under the No BCA-regime.

Proposition 2 - The Effects of Taxes on Production and Consumption Patterns

Suppose a coalition S with m signatories has formed in the first stage, let $1 < m < n$, and let $t_i > t_j$ under both regimes.

1) *Under the No BCA-regime, a firm sells the same quantities to all markets. Under the BCA-regime, quantities are differentiated. Signatory firms sell more to signatory than to non-signatory markets (except for $\gamma = 0$ in which case quantities for all markets are the same) and non-signatory firms sell more to non-signatory than to signatory markets.*

2) Under the No BCA-regime, in every market, non-signatory firms sell more than signatory firms. Under the BCA-regime this is also true, except in signatory markets where all firms sell the same quantities.

3) Under the No BCA-regime, consumption in all markets is the same. Under the BCA-regime, consumption in every signatory market is lower than in every non-signatory market.

Part 1 in Proposition 2 stresses that markets do not matter for firms' output decisions under the No BCA-regime. In contrast, under the BCA-regime, markets matter because effective taxes faced by firms for different markets differ. For firms located in signatory countries, signatory markets are more attractive because their competitors face tax t_i too whereas in non-signatory markets they face t_i but their competitors only face t_j, with $t_i > t_j$. For firms located in non-signatory countries, the mirror image argument applies. Hence, they prefer to sell more to non-signatory than signatory markets.

Part 2 in Proposition 2 has implications for the profits of firms. Because taxes in signatory countries are higher than in non-signatory countries by assumption, firms located in non-signatory countries sell more in each market and hence will have higher profits than firms located in signatory countries. This advantage of non-signatory over signatory firms is partially offset through BCAs because quantities and hence profits in signatory markets are now the same (though quantities and profits in non-signatory markets are still higher). Thus, taken together, BCAs allow to reduce the difference in profits between signatory and non-signatory firms but do not eliminate this difference.

It is noteworthy that the conclusion emerging from Part 2 in Proposition 2 does not change when moving from a firm's perspective to a country's welfare perspective. Signatory firms' profits are lower than non-signatory firms' profits for two reasons: lower output and higher taxes. However, from a country's welfare perspective, taxes are welfare neutral in this simple model: firms' tax bills are revenues of governments. Thus, what matters at the country level are profits excluding tax payments, i.e., gross profits. Hence, also from a country's welfare perspective, signatory countries are disadvantaged because of lower gross profits, which is a result of lower output.

Part 3 in Proposition 2 stresses that consumers in signatory and non-signatory countries consume exactly the same individual and total quantities under the No BCA-regime. As a result, the consumer surplus in signatory and non-signatory countries is the same. In contrast, under the BCA-regime, the quantity of each variety consumed is lower in a signatory market than in a non-signatory market. This is because all quantities supplied to signatories' markets face the high tax t_i whereas those supplied to non-signatories' markets depend on a mix of high t_i and low t_j. Consequently, also the consumer surplus is lower in signatory countries than in non-signatory countries. Thus, it is important to note that border carbon adjustments in the form of import tariffs do not improve but rather they worsen the situation for consumers compared to the No BCA-regime.

In order to complete our understanding of the difference between the two regimes, we note two more items beyond Proposition 2.

First, in terms of environmental damages, even though total production and hence global pollution will be reduced when moving from the No BCA- to the BCA-regime (see Result 1 in section 4 below), in relative terms, nothing changes between signatory and non-signatory countries, as damages are the same for both groups of countries.

Second, in terms of revenues, signatories obtain revenues from the import tariff under the BCA-regime, which is not available under the No BCA-regime. Those revenues are de facto a transfer from firms located in non-signatory countries to signatory governments and hence constitute a loss to non-signatory governments.

As a result, overall, using BCAs to support a climate agreement implies reducing (though not eliminating) the competitive advantage of non-signatory firms over signatory firms, and collecting tariff revenues from foreign firms by signatory governments, but disadvantaging consumers in signatory countries. As our later analysis will illustrate, the overall effect works to the advantage of signatory over non-signatory countries, explaining, among other factors, why large agreements are stable under the BCA-regime, but not under the No BCA-regime.

4. Second Stage

In the second stage, governments choose their taxes. All governments understand how taxes affect both quantities and welfare. That is, governments have solved the third stage of the game, as laid out in the previous section. As pointed out in subsection 2.1, governments which are part of coalition S choose their tax t_i cooperatively by maximizing the aggregate welfare of all coalition members, whereas governments which do not belong to S choose their tax t_j non-cooperatively by maximizing their individual welfare. The simultaneous solution of the m first order conditions of signatory governments and the $n-m$ first order conditions of non-signatory governments delivers equilibrium taxes, which are a function of all parameters of the model. In particular, equilibrium taxes depend on the size of coalition S, m, and hence we may write $t_i^*(m)$ and $t_j^*(m)$.

Under the No BCA-regime, equilibrium taxes are huge terms and they are even bigger under the BCA-regime. Even though we have been able to derive analytical results for the No-BCA (see Finus and Al Khourdajie, 2018), we have not been able to do the same for the BCA-regime. Hence, as we want to compare regimes, we need to resort to simulations.[10] In Appendix 2, we provide a detailed description of how we conducted simulations and how we choose parameters in order to ensure a comprehensive coverage of the parameter space and robust results. In the following,

[10]Simulations are common in the literature on international environmental agreements (IEAs; see, e.g., the collection of articles in Finus and Caparrós, 2015). Even in simpler models, the determinination of stable agreements frequently relies on simulations.

we are mainly interested in qualitative conclusions and we focus less on detailed quantative results.

For the interpretation of the subsequent results, we note that we we assume $n = 10$ countries and consider three values of the TFV-parameter: no TFV with $\gamma = 1$, partial TFV with $\gamma = 0.5$, and full TFV with $\gamma = 0$. Furthermore, as explained in Appendix 2, for the parameters a, c, and δ their ratio matters but their absolute values do not matter. This gives rise to six values for parameter a, which we denote by $a_1(\gamma, \delta)$,..., $a_6(\gamma, \delta)$. Loosely speaking, a_k is the ratio between $a - c$, which can be interpreted as the market size corrected for production costs, i.e., a measure of the net benefits from production and consumption, and δ the damage evaluation. Hence, the smaller the value a_k, the smaller are the net benefits from production and consumption compared to the environmental damages that are associated with these activities. It is important to note for the following discussion that all parameter ranges imply $t_i^*(m) \geq t_j^*(m)$, $1 < m < n$, under both regimes.

Result 1 - Comparing Equilibrium Taxes Across Regimes

Denote equilibrium taxes under both regimes with superscripts No BCA and BCA, respectively, and assume $1 < m < n$.

a) *Under the BCA-regime signatories' equilibrium taxes are higher than under the No BCA-regime:* $t_i^{*BCA} > t_i^{*No\ BCA}$ *for all* m.

b) *Under the BCA-regime non-signatories' equilibrium taxes are higher than under the No BCA-regime for* $\gamma = \{0, 0.5\}$*:* $t_j^{*BCA} > t_j^{*No\ BCA}$ *for all* m*. For* $\gamma = 1$*, they can be higher,* $t_j^{*BCA} > t_j^{*No\ BCA}$*, or lower,* $t_j^{*BCA} < t_j^{*No\ BCA}$.

c) *Under the BCA-regime, total output is lower and hence total emissions are lower than under the No BCA-regime for all* m.

BCAs provide signatory governments with an additional strategic tool to internalize externalities from emissions but also to protect their firms' competitiveness in signatory markets. Furthermore, importantly, it also serves as an additional source of revenues obtained from import tariffs. Therefore, taxes of signatory governments are higher under the BCA-regime than under the No BCA-regime (Result 1,a).

From the perspective of non-signatory governments, BCAs have the following implications. First, their consumers need to pay higher prices for varieties imported from signatory countries. Second, their firms face an additional tax burden at the border with signatory markets, which will negatively affect their profits. Third, they face a loss of potential tax revenue, i.e., the tax revenue generated by their firms, of which some portion goes into signatory governments' coffers. Therefore, the reaction of non-signatory governments is complex. On the one hand, non-signatory governments could raise their taxes in order to protect their tax revenues. On the other hand, they could lower their taxes to protect their consumers and to strenghten the competitivness of their firms, at least in their own markets. According to Result

1,b, for sufficiently low values of γ, the tax revenue effect always dominates the consumer and producer protection effect. With low values of γ, firms face little competition and consumers love variety, so there is not much need for protection and hence $t_j^{*BCA} > t_j^{*No\ BCA}$.

However, even if $t_j^{*BCA} < t_j^{*No\ BCA}$, the overall tax level under the BCA-regime is higher than under the No-BCA-regime such that total output and hence total emissions are always lower for every coalition of size m, as Result 1,c highlights.

5. First Stage

5.1. Preliminaries

In this section, we derive the size of stable coalitions, m^*. In order to evaluate outcomes, we consider a relative welfare measure proposed in Eyckmans and Finus (2006), called the closing the gap index (CGI), which is defined as follows:

$$CGI(m^*) := \frac{\sum_{k\in N} W_k(m^*) - \sum_{k\in N} W_k(m=1)}{\sum_{k\in N} W_k(m=n) - \sum_{k\in N} W_k(m=1)} \bullet 100\,. \tag{11}$$

This index measures to what extent a stable agreement with m^* members closes the gap between the grand coalition, a coalition including all countries ($m = n$), corresponding to the social optimum, and no agreement ($m = 1$), corresponding to the non-cooperative equilibrium, the classical Nash equilibrium without coalition formation. The CGI expresses this in percentage terms. Hence, for instance, if the grand coalition is stable, the CGI is 100% whereas if no agreement is stable, i.e., no non-trivial coalition with at least two members is stable, the CGI is 0%.[11] All details of the results and properties, which are discussed below, are provided in Appendix 3 in our working paper.

5.2. Open Membership: Simultaneous Coalition Formation Process

For our standard assumption of open membership and a simultaneous coalition formation process, we obtain the following result.

Result 2 - Equilibrium Coalitions under Open Membership

Let m^* denote the size of a stable coalition under open membership and let $CGI(m^*)$ denote the associated closing the gap index of stable agreements as defined in Eq. (11). Then, under the No BCA- and BCA-regime, we find the result displayed in Table 1.

[11] In principle, the nominator could be negative for some m^*, $1 < m^* < n$, and hence also the CGI. In these particular cases, it seems sensible to indicate only that the CGI is negative but not to provide exact numbers. See Result 3 below.

Table 1 to be inserted here

Result 2 supports two conclusions.

First, under the No BCA-regime, stable agreements are small if they exist at all. Only if the value of the taste for variety parameter γ is sufficiently small will an agreement with at least two members be stable. But even for $\gamma = 0$, only an agreement with three countries is stable in our model. Accordingly, as the size of stable coalitions m^* is much smaller than the total number of countries, n, the closing the gap index is at best small if not zero. In our model, the strategic interaction among markets and hence governments is related to the TFV-parameter γ (see Propositions 1 and 2 in section 3). The larger the value of γ, i.e., the lower the taste for variety, the stronger will be the strategic interaction among firms and, hence, also among governments. Under the No BCA-regime, strong strategic interaction means strong free-rider incentives and leakage effects.

In a game-theoretic sense, the coalition formation game under the No BCA-regime can be viewed as an n-player (symmetric) chicken game (in pure strategies) with either no or only a small number of players joining an agreement. Non-signatories are always better off than signatories, $W_{i \in S}(m) < W_{j \notin S}(m)$ for every m, $1 < m < n$, as all countries have the same welfare function (ex-ante symmetric players), enjoy the same consumer surplus and suffer the same damages, but gross profits are lower in signatory countries due to the lower outputs of their firms.

Moreover, the incentive to abstain from the agreement shows up in an increase of welfare of non-signatories whenever the agreement grows by one more member, i.e., $W_{j \notin S}(m+1) > W_{j \notin S}(m)$, a property which we call positive external spillovers. Under the No BCA-regime, this property holds for every m, $1 \leq m < n-1$, and all parameter values. The economics behind this property is that whenever the coalition is enlarged, in equilibrium, signatories decrease their outputs whereas non-signatories increase their outputs, though the total output (and hence environmental damages) decreases. Hence, non-signatories benefit from lower damages, higher prices for their goods and a relocation of production to their countries if participation in an agreement increases. Moreover, carbon leakage makes it attractive to abstain from an agreement.

The "small coalition paradox" under the No BCA-regime due to strong free-rider incentives is particularly miserable as global welfare constantly increases with membership, a property called full cohesiveness. Thus, the CGI-values are small.

Second, in contrast, under the BCA-regime, large coalitions are stable, including the grand coalition, with accordingly large values of the closing the gap index. BCAs put signatories in a strong position, which is particularly useful for them if there is a high degree of strategic interaction between signatories and non-signatories. According to Result 2, the grand coalition is stable for $\gamma = 0.5$ and $\gamma = 1$. Hence, the CGI is 100%. For $\gamma = 0$, we find stable coalitions below full participation. However, the CGI is still substantially above 85%.

In a game-theoretic sense, the coalition formation game under the BCA-regime can no longer be viewed as a chicken game as now $W_{i \in S}(m) > W_{j \notin S}(m)$ in most cases. BCAs improve the welfare of signatories and mostly lower the welfare of non-signatories compared to the No BCA-regime. Hence, it is more likely that internal stability ($W_{i\in S}(m) \geq W_{j\notin S}(m-1)$) holds not only for small but also for large agreements m. Moreover, coalition formation may no longer be associated with positive external spillovers but are associated with negative external spillovers (i.e., $W_{j\notin S}(m+1) < W_{j\notin S}(m)$). That is, not being a member of an environmentally motivated customs union is not rewarded but is rather increasingly punished if an agreement grows.

In our model, negative external spillovers are favored by large values of γ and a_k under the BCA-regime. High values of γ indicate a strong strategic interaction of markets. Hence, the larger the value of γ, the more are non-signatories disadvantaged at the expense of signatories. High values of a_k indicate a relatively large market size compared to the valuation of environmental damages. Consequently, the larger the value a_k the more will non-signatories lose from the extraction of tax revenues through import tariffs and the less they will benefit from a reduction of damages if an agreement grows.

Altogether, BCAs are a game changer, making it less attractive to abstain from and more attractive to join an agreement. However, not only in terms of membership, but also in terms of global welfare, agreements supported by BCAs perform well. Even in the worst case in which large though not full participation is achieved, the gap between full and no cooperation is closed to a large extent. Interestingly, a high degree of strategic interaction and linkage between markets (captured by large values of the TFV-parameter γ in our model), which is detrimental to participation under the No BCA-regime, is conducive to participation under the BCA-regime.

5.3. Alternative Coalition Formation Processes

In this subsection, we analyze the robustness of our conclusions regarding our assumptions. Notwithstanding this technical point, more importantly, we are interested in what would change if agreements are not open to the accession of outsiders (open membership) but are subject to an approval process by current members (exclusive membership). Therefore, we consider two alternative features: a) a sequential instead of a simultaneous coalition formation process and b) exclusive instead of open membership.

5.3.1 Open Membership: Sequential Coalition Formation

The standard assumption in the literature on coalition formation is a simultaneous choice of membership as considered so far (see Bloch, 2003; Yi 1997). However, in reality, agreements typically form sequentially, with some initiators moving first and some laggards joining later. The simplest extension of our game could assume that two countries form a coalition in a first instance. Subsequently, others may join if they benefit from accession.

In considering such a sequential process, any intermediate coalition of size m must be - first of all - internally stable. Outsiders will accede to an intermediate coalition of size m if this is beneficial to them. In other words, such a coalition is not externally stable. (See the definition of internal and external stability in section 2.1.) In the case of symmetric players, if a coalition of size m is externally unstable, then this implies that a coalition of size $m+1$ is strictly internally stable, i.e., $W_{i\in S}(m) - W_{i\notin S}(m-1) > 0$ (Carraro and Sinsicalco, 1993). Thus, if there is a sequence of coalitions m in the interval $1 < m \leq \overline{m} \leq n$ which are all strictly internally stable, then there will be a sequence of accessions until $\overline{m}$ is reached. Moreover, $\overline{m}$ is externally stable and hence $\overline{m} = m^*$. The reason is simple: either a) $\overline{m} = n$, in which case $\overline{m}$ is externally stable as no outsider is left or b) $\overline{m} < n$, but then $m = \overline{m} + 1$ is internally unstable (by the initial assumption that only coalitions up to $m = \overline{m}$ are internally stable) and hence $\overline{m}$ is externally stable.

Therefore, the only question that remains is whether there exists such a sequence of strictly internally stable coalitions and if so, whether the final coalition emerging from such a sequence of accessions, $\overline{m}$, is smaller or identical to the size of the stable coalition m^*, as reported for the simultaneous coalition formation process above. Since the answer to the first part of the question is affirmative and the answer to the second part reveals $\overline{m}= m^*$, all results displayed for a simultaneous coalition formation process in Result 2 are exactly the same for a sequential coalition formation process. Thus, our results are robust regarding this modification of assumption.

5.3.2 Exclusive Membership: Sequential Coalition Formation

We now consider exclusive membership. It is noteworthy that although international environmental agreements are typical of the open membership type, trade agreements classically feature exclusive membership, including accession to the WTO or the European Common Market. That is, most trade agreements only accept accession by way of the approval of current members.

In principle, exclusive membership can be considered for a simultaneous and a sequential coalition formation process. However, for the sake of expositional simplicity, we consider only the more interesting case of a sequential process, referring the reader to our working paper for the alternative assumption of a simultaneous formation process.

For a sequential process under open membership, we have argued above that there exists a sequence of accessions such that finally $\overline{m} = m^*$ is reached. That is, all intermediate coalitions are internally stable and outsiders have an incentive to join such coalitions. Now, under exclusive membership, we need to ask whether coalition members of an intermediate coalition would have a reason to deny the accession of outsiders. They will have no reason if the expansion increases their welfare, i.e., $W_{i\in S}(m+1) > W_{i\in S}(m)$, which we call positive internal spillovers, but will refuse accession if the opposite is true, namely, $W_{i\in S}(m+1) < W_{i\in S}(m)$, which we will call negative internal spillovers.

Thus, if we denote the size of the stable coalition under exclusive membership by m^{**}, recalling that we denoted stable coalitions under open membership by m^*, then $m^{**} \leq m^*$ must generally be true. Hence, the interesting question is whether and under which conditions this inequality is strict, i.e., $m^{**} < m^*$ and what are the implications for global welfare. This is shown in Result 3.

Result 3 - Equilibrium Coalitions under Exclusive Membership

Assume exclusive membership and a sequential coalition formation process. Let m^{**} denote the size of a stable coalition under the No BCA- and BCA-regime. Then, we find the results displayed in Table 2:

Table 2 to be inserted here

Comparing Results 2 and 3, it is evident that exclusive membership makes no difference for the No BCA-regime, $m^{**} = m^*$. This is also true for the BCA-regime for $\gamma = 0$. However, for the other two TFV-parameter values, $\gamma = 0.5$ and $\gamma = 1$, stable coalitions are smaller than under open membership, i.e., $m^{**} < m^*$, with associated CGI-values which are also lower and even negative in some cases. Negative CGI-values imply that global welfare for m^{**} is even below the level without agreement. For $\gamma = 0.5$ and $\gamma = 1$, we observe that membership is strongly reduced compared to open membership and GGI-values are particulary low and negative for large values of a_k.

Hence, in the following, we need to address two questions. First, why and under which conditions are there negative internal spillovers under the BCA-regime which lead to $m^{**} < m^*$? Second, why and under which conditions are these agreements associated with low CGI-values?

Under the BCA-regime, signatories enjoy a strong position towards non-signatories due to import tariffs. This strength is growing at the beginning of the expansion of an agreement: additional members mean lower damages and lower leakage effects, which can be controlled by import tariffs. As a result, at the beginning, an expansion is always associated with positive internal spillovers. However, at some level $\tilde{m} < n$, these internal spillovers may become negative, in which case the expansion stops and $m^{**} = \tilde{m}$. A representative example, which illustrates $m^{**} < m^*$ under the BCA-regime, is drawn in Figure 1a. $W_{i \in S}(m)$ reaches its peak at $m = \tilde{m} = 5$. Hence, $m^{**} = 5$ under exclusive membership whereas under open membership $m^* = 10$ (see Results 2 and 3)

Figure 1 about here

In our model, negative internal spillovers under the BCA-regime above a threshold coalition size $\tilde{m}$ are favored by high values of γ and a_k. Note that these are the same conditions which favor negative external spillovers. Hence, the reasons are very similar. The larger the value of γ, the stronger is the advantage of signatories over non-signatories. The larger the value of a_k, the more attractive it is for signatories to extract tax revenues through import tariffs from non-signatories compared to the motive to reduce environmental damages. With growing membership, the basis to extract tax revenues decreases. Consequently, above some level of participation $\tilde{m} < n$, the marginal decrease in tariff revenues exceeds the marginal benefit of damage reduction from an increase of membership.

The reason for low CGI-values for those coalitions with small membership under the BCA-regime relates to the failure of full cohesiveness. We may recall that full cohesiveness always holds, i.e., global welfare continuously increases with membership, under the No-BCA regime. Under the BCA-regime, this may no longer be true as Figure 1b illustrates for our representative example. Already at $m = 3$, global welfare starts to decline with an enlargement of the coalition until reaching the lowest global welfare at $m = 8$ and then picking up again until the grand coalition is reached.[12] As $m^{**} = 5$ in this example, global welfare is even below the level without agreement. Hence, the CGI-value is negative, as reported in Result 3.

Declining segments of global welfare as a function of membership are due to strong negative external spillovers. That is, the negative impact of the expansion of an agreement on non-signatories is stronger than any positive effect on signatories. Consequently, in our model, declining segments of global welfare are favored by high values of γ and a_k. As discussed above, these are also the conditions which favor a restriction of membership through signatories under exlusive membership. Therefore, the low and even negative CGI-values for some parameter values under the BCA-regime are the result of two features: a) agreements which fall substantially short of full participation and b) agreements without full membership which are associated with low global welfare.

Viewed together, under the BCA-regime, those conditions, which negatively impact on global welfare as long as participation is not close to full participation, are also the conditions which reduce the free-rider incentive while they also encourage the restriction of membership by signatories. This trade-off does not materialize under open membership as stable agreements include all or almost all countries. That is, punishment of non-participation through import tariffs remains a threat but is normally not implemented. In contrast, under exclusive membership, individual rationality may impede global rationality. Individual rationality may encourage signatories to restrict membership and therefore BCAs would be implemented, which may be detrimental to global welfare.

Overall, this allows us to draw at least two important policy conclusions. First, the general fear by most economists that any restriction on trade may cause large welfare losses, even if motivated by environmental concerns, is supported by our model. BCAs may be associated with a global welfare loss if full cooperation is not achieved. Second, nevertheless, border carbon adjustments can be a useful tool to enforce climate agreements and to increase global welfare. However, membership should not be restricted in those agreements in order to avoid the fact that some

[12] As the grand coalition generates the highest global welfare, it is clear that if there is a decreasing segment of global welfare between $m = 1$ and $m = n$, there must also be an increasing segment.

governments could potentially hijack this instrument for their own interest. BCAs are designed to internalize a global externality. Therefore, climate agreements supported by BCAs should remain open to all countries that would like to join for the global good.

6. Summary and Conclusion

In a stylized intra-industry trade model with horizontal product differentiation and taste for variety (TFV) by consumers we studied the formation, stability and success of international climate agreements. In the tradition of the game-theoretic literature on international environmental agreements (IEAs), we modeled a three-stage game in which governments first decide whether to join a climate agreement, then they decide on their policy levels and finally firms choose their outputs. The model captures the strategic interaction between signatories and non-signatories with an endogenous choice of strategies in each stage. We considered two regimes. Under the first regime, governments have only an emission tax at their avail to correct externalities. Under the second regime, signatories to a climate agreement have an additional policy tool, namely an import tariff. The tariffs, the border carbon adjustments (BCAs), are chosen such that the effective tax on imports is the same as the tax on domestic production. We labeled the first regime the No BCA-regime and the second the BCA-regime. Export rebates were not part of our BCA-regime as they may be difficult to justify within WTO-rules, may invite retaliatory measures by non-signatories but also because import tariffs are already able to deliver what most scholars intuitively suspect, namely, they enforce larger stable and effective climate agreements.

We showed that BCAs are a game changer in several respects. Firms located in signatory countries, facing higher taxes than their rivals in non-signatory countries, now play on equal terms with their rivals, at least in their home market. Thus, the difference in profits is reduced through BCAs, though it does not vanish. Signatory governments also benefit from tariff revenues which, at the same time, constitute a loss of tax revenues to non-signatory governments. However, not all effects work to the absolute or relative advantage of signatory countries. With BCAs consumers in signatory countries are disadvantaged as they face higher prices than consumers in non-signatory countries. Nevertheless, at the aggregate welfare level of countries, the first two effects dominate the third effect. That is, signatories' welfare increases under the BCA-regime compared to the No BCA-regime and, in most cases, the reverse is true for non-signatory countries. Thus, joining an agreement becomes more attractive under the BCA-regime rather than under the No-BCA regime.

A variation of this theme showed up in the properties of the coalition formation game.

Under the No-BCA regime, non-signatories' welfare increases with the expansion of the agreement, which we called positive external spillovers. This is a result of the non-excludability of public good provision. Larger climate agreements imply more ambitious emission reduction targets by signatories from which non-signatories benefit at no cost. Moreover, in the context of trade, non-signatories benefit from a relocation of production to their countries. These leakage effects also render agreements without full participation not very effective and discourage participation in a climate agreement. All this explains why only small climate agreements are stable under the No BCA-regime, which, at best, only marginally close the gap between no and full cooperation in terms of global welfare.

Under the BCA-regime, large agreements can be stable, including an agreement comprising all countries. Import tariffs reduce the leakage effect, provide a benefit to signatory countries and may even impose negative external spillovers on non-signatories, making it attractive for outsiders to join a climate agreement. Hence, under open membership, if accession to an agreement is not restricted, stable agreements comprise all or almost all countries and are associated with large global welfare gains. That is, BCAs work like a threat, but, in equilibrium, they are either not implemented or they are only implemented at a small scale. However, under exclusive membership, only partial agreements may emerge. We could show that in those cases signatories may find it individually rational to restrict membership because with increasing membership the basis to extract tariff revenues decreases. However, in these cases, this is not in line with global rationality: global welfare may be low.

Thus, overall, climate agreements aiming at internalizing a global environmental externality, benefit from using BCAs in order to support this goal. However, they should be of the open membership type in the tradition of environmental agreements and not of the exclusive membership type as this is true for most trade agreements.

Our analysis made some simplifying assumptions. In the following, we briefly discuss the possible implications of relaxing these assumptions.

First, we did not consider export rebates for the reasons given above. Similar arguments would apply to trade bans, which would also be more difficult to justify under WTO-rules than import tariffs and may also provide more reasons for non-signatories to retaliate (see the last point below). With export rebates and trade bans, signatories would gain an even stronger position towards non-signatories which, as we have shown, is not necessary to make BCAs successful under open membership. We expect that the possible negative impacts of BCAs under exclusive membership would be more pronounced.

Second, we could allow for asymmetric welfare functions of countries (Nkuija, 2003). Though more realistic, the driving forces/properties identified in our model with ex-ante symmetric players would not disappear. If tariff and tax revenues among signatories were used to compensate those benefiting from cooperation less than others, asymmetry does not have to be an obstacle and it may even be an asset for cooperation, as this has already been demonstrated to emerge from asymmetric IEA models (without trade) for instance by Finus and McGinty (2019) and Weikard (2009).

Finally, one could speculate what happens if non-signatories would retaliate on

BCAs.[13] As pointed out above, the justification for BCAs is not strategic in that countries want to improve their terms of trade through tariffs. They only correct a global environmental externality and are typically only used as a threat if not all countries join a climate agreement. Thus, if successful, they are normally not implemented. Nevertheless, one could imagine that in those cases in which the grand coalition is not stable, non-signatories could retaliate, even though this would not be in accordance with WTO-rules. Certainly, retaliatory BCAs are not possible because, per assumption, taxes in signatory countries are higher than in non-signatory countries. Thus, non-signatories could only impose an import tariff that has no environmental justification. If signatories adhered to the BCA-rules, this would weaken their position and as a result would weaken the positive impact of BCAs which we have identified. However, a reaction to such an optimal retaliatory tariff by non-signatories could also imply that signatories give up BCA-rules (i.e., WTO-rules): they chould choose an optimal tariff without the restriction that taxes of non-signatories plus tariff cannot exceed domestic taxes. Hence, we would end up in an optimal tariff game, without any reference to BCAs. Predictions would require a careful modeling of such a tariff game, which would certainly warrant a paper in its own right. Our conjecture would be that such a tariff game could also establish large stable coalitions in some cases. However, whenever the grand coalition is not stable, the negative global welfare implications would be even more severe than under our BCA-regime.

Acknowledgments

We would like to acknowledge the helpful comments by Sergio Currarini, Lucy O'Shea, Javier Rivas, Santiago Rubio, Aart de Zeeuw, two anonymous reviewers as well as the guest editors, Cameron Heburn, Nicholas Stern and Joseph Stiglitz. Alaa Al Khourdajie would like to thank Pedro Pintassilgo for his hospitality during a research visit at the University of the Algarve, Faro, Portugal in 2016 and the many fruitful discussions. The authors bear the responsibility of any errors and omissions that may remain.

7. References

1. Al Khourdajie, A. and M. Finus (2018), International Environmental Agreements and Carbon Border Adjustments. Bath Economic Research Papers 71/18, University of Bath.

2. Anoulies L. (2014), The Strategic and Effective Dimensions of Border Tax Adjustment. "Journal of Public Economic Theory", vol. 17(6), pp. 824-847.

[13] We owe this interesting point to an anonymous reviewer of our paper.

3. d'Aspremont, C., A. Jacquemin, J.J. Gabszewicz and J.A. Weymark (1983), On the Stability of Collusive Price Leadership. "Canadian Journal of Economics", vol. 16(1), pp. 17-25.

4. Baksi, S. and A.R. Chaudhuri (2009), On Trade Liberalization and Transboundary Pollution. "Economics Bulletin", vol. 29(4), pp. 2605-2612.

5. Baksi, S. and A.R. Chaudhuri (2014), Trade, Border Tax Adjustments, and the Stability of International Environmental Agreements. Working Paper.

6. Baksi, S. and A.R. Chaudhuri (2017), International Trade and Environmental Cooperation among Heterogeneous Countries. In: O. Kayalca, S. Cagatay and H. Mihci (eds.), Economics of International Environmental Agreements: A Critical Approach, Routledge, pp. 97-115.

7. Barrett, S. (1994a), Self-Enforcing International Environmental Agreements. "Oxford Economic Papers", vol. 46, pp. 878-894.

8. Barrett, S. (1994b), Strategic Environmental Policy and International Trade. "Journal or Public Economics", vol. 54, pp. 325-338.

9. Barrett, S. (1997). The Strategy of Trade Sanctions in International Environmental Agreements. "Resource and Energy Economics", vol. 19(4), pp. 345-361.

10. Benchekroun, H., and N. van Long (1998), Efficiency Inducing Taxation for Polluting Oligopolists. "Journal of Public Economics", vol. 70(2), pp. 325-342.

11. Bloch, F. (2003), Non-Cooperative Models of Coalition Formation in Games with Spillovers. In: Carraro, C. (ed.), The Endogenous Formation of Economic Coalitions. Edward Elgar, Cheltenham, UK et al., ch. 2, pp. 35-79.

12. Böhringer, C, B. Bye, T. Fæhn, and K. E. Rosendahl (2017a), Output-based Rebating of Carbon Taxes in the Neighbors Backyard. "Canadian Journal of Economics", vol. 50(2), pp. 426-455.

13. Böhringer, C. A. Lange, and T. F. Rutherford (2014), Optimal Emission Pricing in the Presence of International Spillovers: Decomposing Leakage and Terms-of-trade Motives. "Journal of Public Economics", vol. 110, pp. 101-111.

14. Böhringer, C., A. Müller, and J. Schneider (2015), Carbon Tariffs Revisited. "Journal of the Association of Environmental and Resource Economists", vol. 2(4), pp. 629-672, 2015.

15. Böhringer, C. K. E. Rosendahl and H. B. Storrøstenc (2017b), Robust Policies to Mitigate Carbon Leakage. "Journal of Public Economics", vol. 149, pp. 35-46.

16. Brander, J.A. and B.J. Spencer (1985), Export Subsidies and International Market Share Rivalry. “Journal of International Economics”, vol. 18, pp. 83-100.

17. Branger, F. and & P. Quirion (2014), Would Border Carbon Adjustments Prevent Carbon Leakage and Heavy Industry Competitiveness Losses? Insights from a Meta-analysis of Recent Economic Studies. “Ecological Economics”, vol. 99(C), pp. 29-39.

18. Burguet, R. and L. Sempere (2003), Trade Liberalization, Environmental Policy, and Welfare. “Journal of Environmental Economics and Management”, vol. 46(1), pp. 25-37.

19. Caron, J. (2012), Estimating Carbon Leakage and the Efficiency of Border Adjustments in General Equilibrium — Does Sectoral Aggregation Matter? “Energy Economics”, vol. 34, pp. 111-126.

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

21. Conrad, K. (1993). Taxes and Subsidies for Pollution-Intensive Industries as Trade Policy. “Journal of Environmental Economics and Management” 25, pp. 121–135.

22. Copeland, B. R. and M. S. Taylor (1995), Trade and Transboundary Pollution. “American Economic Review”, vol. 84(4), pp. 851-874.

23. Di Comite, F., J-F. Thisse and H. Vandenbussche (2014), Verti-zontal Differentiation in Export Markets. “Journal of International Economics”, vol. 93(1), pp. 50-66.

24. Dixit, A. K. and J. Stiglitz (1977), Monopolistic Competition and Optimum Product Diversity. “The American Economic Review”, vol. 67(3), pp. 297-308.

25. Eichner, T. and R. Pethig (2013), Self-Enforcing Environmental Agreements and International Trade. “Journal of Public Economics”, vol. 102, pp. 37-50.

26. Eichner, T. and R. Pethig (2015), Self-Enforcing International Environmental Agreements and Trade: Taxes Versus Caps. Oxford Economic Papers vol. 67, pp. 897-917.

27. Eyckmans, J. and M. Finus (2006), New Roads to International Environmental Agreements: The Case of Global Warming. “Environmental Economics and Policy Studies”, vol. 7, pp. 391-414.

28. Eyland, T. and G. Zaccour (2012), Strategic Effects of a Border Tax Adjustment. “International Game Theory Review”, vol. 14(3), pp. 1-22.

29. Eyland, T. and G. Zaccour (2014), Carbon Tariffs and Cooperative Outcomes. "Energy Policy", vol. 65, pp. 718-728.

30. Financial Times (Feb. 12, 2017), A Carbon Tax is the Best Answer to Climate Change. Lakshmi Mittal.

31. Finus, M. and A. Al Khourdajie (2018), Strategic Environmental Policy, International Trade and Self-enforcing Agreements: The Role of Consumers' Taste for Variety. "Strategic Behavior and the Environment", vol. 7, pp. 317-350.

32. Finus, M. and A. Caparrós (2015), Game Theory and International Environmental Cooperation: Essential Readings. Edward Elgar, Cheltenham.

33. Finus, M. and M. McGinty (2019), The Anti-Paradox of Cooperation: Diversity May Pay. Journal of Economic Behavior and Organization, vol. 157, pp. 541-559.

34. Fischer C. and A.K. Fox (2012), Comparing Policies to Combat Emissions Leakage: Border Carbon Adjustments versus Rebates. "Journal of Environmental Economics and Management", vol. 64, pp.199–216.

35. Helm, D., C. Hepburn and G. Ruta (2012), Trade, Climate Change, and the Political Game Theory of Border Carbon Adjustments. "Oxford Review of Economic Policy", vol. 2, pp. 368-394.

36. Helpman, E. and P.R. Krugman (1985), Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy. The MIT Press. Cambridge, Massachusetts. London, UK.

37. Hoel, M. (1996), Should a Carbon Tax be Differentiated Across Sectors? "Journal of Public Economics", vol. 59, pp. 17-32.

38. Irfanoglu, Z.B., J.P. Sesmero, and A. Golub (2015), Potential of Border Tax Adjustments to Deter Free Riding in International Climate Agreements. "Environmental Research Letters", vol. 10(2), article no. 024009.

39. Keen, M. and C. Kotsogiannis (2014), Coordinating Climate and Trade Policies: Pareto Efficiency and the Role of Border Tax Adjustments. "Journal of International Economics", vol. 94(1), pp. 119-128.

40. Kennedy, P.W. (1994), Equilibrium Pollution Taxes in Open Economies with Imperfect Competition. "Journal of Environmental Economics and Management", vol. 27, pp. 49-63.

41. Loke, W-H. and L.A. Winters (2012), Why is Open Regionalism Not Always Good? "Journal of Economic Integration", vol. 27(3), pp. 437-465.

42. Markusen, J.R. (1975), International Externalities and Optimal Tax Structures. “Journal of International Economics”, vol. 5, pp. 15-29.

43. Mehling, M., H. van Asselt, K. Das, S. Dröge and C. Verkuijl (2017), Designing Border Carbon Adjustments for Enhanced Climate Action. “American Journal of International Law”, vol. 113(3), pp. 433-481.

44. Nkuiya, B. (2013), Trade Structure, Transboundary Pollution and Multilateral Trade Liberalization: the Effects on Environmental Taxes and Welfare. “Environmental and Resource Economics”, vol. 55(3), pp. 337-355.

45. Nordhaus, W. (2015), Climate Clubs: Overcoming Free-riding in International Climate Policy. “American Economic Review”, vol. 105(4), pp. 1339-1370.

46. Stiglitz, J.E. (2006), A New Agenda for Global Warming. Economists’ Voice. July 2016,, 3(7). https://www.carbontax.org/wp-content/uploads/2007/09/stiglitz-_-a-new-agenda-for-global-warming-_-economists-voice-_-july-2006.pdf

47. Tsakiri, N., M. S. Michael and P. Hatzipanayotou (2014), Asymmetric Tax Policy Responses in Large Economies With Cross-Border Pollution. “Environmental and Resource Economics”, vol. 58, pp. 563-578.

48. Ulph, A. (1996), Environmental Policy and International Trade when Governments and Producers Act Strategically. “Journal of Environmental Economics and Management”, vol. 30, pp. 265-281.

49. Vlassis, N. (2013), The Welfare Consequences of Pollution-Tax Harmonization. “Environmental and Resource Economics”, vol. 56, pp. 227-238.

50. Weikard, H.-P. (2009), Cartel Stability under an Optimal Sharing Rule. “Manchester School”, vol. 77(5), pp. 575-593.

51. Weitzel, M., M. Hübler and S. Peterson (2012), Fair, Optimal or Detrimental? Environmental vs. Strategic Use of Border Carbon Adjustment. “Energy Economics”, vol. 34, pp. 198-207.

52. Yi, S-S. (1996), Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism Is Good. “Journal of International Economics, vol. 41, pp. 153-177.

53. Yi, S.-S. (1997), Stable Coalition Structures with Externalities. “Games and Economic Behavior”, vol. 20, pp. 201-237.

54. Yi, S-S. (2000), Free Trade Areas and Welfare: An Equilibrium Analysis. “Review of International Economics”, vol. 8(2), pp. 336-347.

55. Zenghelis, D. and N. Stern (2009), Principles for a Global Deal for Limiting the Risks from Climate Change. “Environmental and Resource Economics”, vol. 43, pp. 307–311.

Appendix 1

In stage 3, the output vector has the structure displayed in the table below.

	Sign. firm i	Non-sign. firm j	Total Consumption
Sign. market k	q^*_{ki}	q^*_{kj}	$Q^*_{k.}$
Non-sign. market l	q^*_{li}	q^*_{lj}	$Q^*_{l.}$
Total Production	$Q^*_{.i}$	$Q^*_{.j}$	Q

We start by deriving quantities consumed in different markets and then we compute aggregate production levels. In order to cover both regimes, the No-BCA- and the BCA-regime, we display quantities with the adjustment parameter ϕ where appropriate, noting that by setting $\phi = 0$ ($\phi = 1$) quantities in the No-BCA- (BCA-) regime would be obtained. It is important to note that quantities are a function of equilibrium taxes, which are determined in stage 2 and those taxes will differ between the No-BCA- and the BCA-regime. It is also important to point out that for the purpose of the discussion in the text with a focus on signatories and non-signatories, we already make use of the fact that in equilibrium in stage 2, all signatories choose the same tax t_i and all non-signatories choose the same tax t_j.

Non-signatory Markets:

The profit of firm i in market l, being located in a signatory country, is given by $\pi_{li} = q_{li}(p_{li} - c - t_i)$. Substitution of the inverse demand function in Eq. (3) in the text, gives the following first order conditions (after the appropriate changes of notation):

$$\frac{\partial \pi_{li}}{\partial q_{li}} = a - c - t_i - (2-\gamma)q_{li} - \gamma Q_{l.} = 0 \iff q_{li} = \frac{1}{2-\gamma}\left[a - c - t_i - \gamma Q_{l.}\right] \quad \text{(A.1)}$$

where $Q_{l.}$ is the total quantity consumed in market l. The right-hand side expression in Eq. (A.1) is the replacement function of firm i ($q_{li} = R_i(Q_{l.})$ which is strictly downward sloping, except for γ= 0 in which case it is a horizontal line. It is evident that a necessary condition for positive quantities is $a > c$.

The profit of firm j in a non-signatory country in market l is given by $\pi_{lj} = q_{lj}(p_{lj} - c - t_j)$ which leads to the following first order condition:

$$\frac{\partial \pi_{lj}}{\partial q_{lj}} = a - c - t_j - (2-\gamma) q_{lj} - \gamma Q_{l.} = 0 \iff q_{lj} = \frac{1}{2-\gamma} \left[a - c - t_j - \gamma Q_{l.} \right] \quad \text{(A.2)}$$

where the right-hand side expression in Eq. (A.2) is the replacement function of firm j ($q_{lj} = R_j(Q_{l.})$. Summing the m first order conditions in Eq. (A.1) and the $n-m$ first order conditions in Eq. (A.2), we derive the aggregate replacement function $\sum_{i \in N} R_i(Q_{l.})$:

$$\sum_{i \in N} R_i(Q_{l.}) := Q_{l.} = \frac{1}{2-\gamma} \left[n(a-c) - mt_i - (n-m)t_j - n\gamma Q_{l.} \right] . \quad \text{(A.3)}$$

The aggregate replacement function is downward sloping over the entire domain. Hence, the equilibrium is unique. Solving Eq. (A.3) for $Q_{l.}$, gives:

$$Q^*_{l.} = \frac{n(a-c) - m(t^*_i - t^*_j) - nt^*_j}{(n-1)\gamma + 2} . \quad \text{(A.4)}$$

Substituting $Q^*_{l.}$ in Eq. (A.4) into Eqs. (A.1) and (A.2), we derive a signatory firm i's variety produced for a non-signatory l's market:

$$q^*_{li} = \frac{(a-c)(2-\gamma) - \left[\gamma(n-m) + (2-\gamma) \right] t_i + \left[\gamma(n-m) \right] t_j}{((n-1)\gamma + 2)(2-\gamma)} \quad \text{(A.5)}$$

and a non-signatory firm j's variety produced for a non-signatory l's market:

$$q^*_{lj} = \frac{(a-c)(2-\gamma) + \gamma m t_i - \left[2 + \gamma(m-1) \right] t_j}{((n-1)\gamma + 2)(2-\gamma)} . \quad \text{(A.6)}$$

Signatory Markets:

The procedure is similar as explained above. The profit of firm i in market k is given by $\pi_{ki} = q_{ki}(p_{ki} - c - t_i)$ which leads to the following first order condition:

$$\frac{\partial \pi_{ki}}{\partial q_{ki}} = a - c - t_i - (2-\gamma) q_{ki} - \gamma Q_{k.} = 0 \iff q_{ki} = \frac{1}{2-\gamma} \left[a - c - t_i - \gamma Q_{k.} \right] . \quad \text{(A.7)}$$

The profit of firm j in market k is given by $\pi_{kj} = q_{kj}(p_{kj} - c - t_j - \phi(t_i - t_j))$, which gives:

$$\frac{\partial \pi_{kj}}{\partial q_{kj}} = a - c - t_j - \phi(t_i - t_j) - (2-\gamma) q_{kj} - \gamma Q_{k.} = 0 \quad \text{(A.8)}$$

$$\iff q_{kj} = \frac{1}{2-\gamma}\left[a - c - t_j - t_j - \phi(t_i - t_j) - \gamma Q_{k.}\right] .$$

Summing the first order conditions in Eqs. (A.7) and (A.8), gives the total equilibrium consumption in a signatory k's market from all varieties:

$$Q^*_{k.} = \frac{n(a-c) - [n - (n-m)(1-\phi)]\, t_i - [(n-m)(1-\phi)]\, t_j}{(n-1)\gamma + 2} . \tag{A.9}$$

Substituting Eq. (A.9) in Eqs. (A.7) and (A.8) above, gives a signatory firm i's variety produced for a signatory k's market:

$$q^*_{ki} = \frac{(a-c)(2-\gamma) - [\gamma(n-m)(1-\phi) + (2-\gamma)]\, t_i + [\gamma(n-m)(1-\phi)]\, t_j}{((n-1)\gamma + 2)(2-\gamma)} \tag{A.10}$$

and a non-signatory firm j's variety produced for a signatory k's market:

$$q^*_{kj} = \frac{(a-c)(2-\gamma) + [\gamma m - \phi(2+\gamma(m-1))]\, t_i - [2+\gamma(m-1) - \phi(2+\gamma(m-1))]\, t_j}{((n-1)\gamma + 2)(2-\gamma)} . \tag{A.11}$$

Production:

Summing Eq. (A.5) over all $n-m$ non-signatory markets and Eq. (A.10) over all m signatory markets gives total production of a signatory firm i's variety to all markets:

$$Q^*_{.i} = \frac{n(a-c)(2-\gamma) - [2n + \gamma n(n-1-m) - \phi\gamma m(n-m)]\, t_i}{((n-1)\gamma + 2)(2-\gamma)} + \frac{[n(n-m) - \phi m(n-m)]\, t_j}{((n-1)\gamma + 2)(2-\gamma)} . \tag{A.12}$$

Similarly, summing Eq. (A.6) over all $n-m$ non-signatory markets and Eq. (A.11) over all m signatory markets gives total production of a non-signatory firm j's variety to all markets:

$$Q^*_{.j} = \frac{n(a-c)(2-\gamma) + [\gamma nm - \phi m(2+\gamma(m-1))]\, t_i}{((n-1)\gamma + 2)(2-\gamma)} - \frac{[\gamma nm + n(2-\gamma) - \phi m(2+\gamma(m-1))]\, t_j}{((n-1)\gamma + 2)(2-\gamma)} . \tag{A.13}$$

Finally, total production/consumption by all countries is given by:

$$Q = \frac{n^2(a-c) - [nm + \phi m(n-m)]\, t_i - [n(n-m) - \phi m(n-m)]\, t_j}{(n-1)\gamma + 2}. \qquad \text{(A.14)}$$

which is Eq. (A.12)+(A.13).

Proposition 1:

Using the quantities derived above, assuming $\phi = 0$ under the No-BCA- and $\phi = 1$ and $t_i > t_j$ under the BCA-regime, we find:

No-BCA-regime: for all quantities of firm i's (j's) variety supplied to every market $h \in N$, $\frac{\partial q^*_{hi}}{\partial t_i} < 0$ ($\frac{\partial q^*_{hj}}{\partial t_j} < 0$), $\frac{\partial q^*_{hi}}{\partial t_j} > 0$ ($\frac{\partial q^*_{hj}}{\partial t_i} > 0$), except for $\gamma = 0$ in which case $\frac{\partial q^*_{hi}}{\partial t_j} = 0$ ($\frac{\partial q^*_{hj}}{\partial t_i} = 0$). BCA-regime: for all quantities supplied to non-signatory markets $l \in N \setminus S$, $\frac{\partial q^*_{li}}{\partial t_i} < 0$ ($\frac{\partial q^*_{lj}}{\partial t_j} < 0$) and $\frac{\partial q^*_{li}}{\partial t_j} > 0$ ($\frac{\partial q^*_{lj}}{\partial t_i} > 0$), except for $\gamma = 0$ in which case $\frac{\partial q^*_{li}}{\partial t_j} = 0$ ($\frac{\partial q^*_{lj}}{\partial t_i} = 0$). For quantities supplied to all signatory markets $k \in S$, $\frac{\partial q^*_{ki}}{\partial t_i} < 0$, $\frac{\partial q^*_{kj}}{\partial t_i} < 0$, $\frac{\partial q^*_{kj}}{\partial t_j} = 0$ and $\frac{\partial q_{ki}}{\partial t_j} = 0$.

Proposition 2:

Using the quantities derived above, assuming $\phi = 0$ under the No-BCA-regime and $\phi = 1$ under the BCA-regime and $t_i > t_j$ under both regimes, we find:

No-BCA-regime: for signatory firm $i's$ outputs (non-signatory firm $j's$ outputs) supplied to all markets $k \in N$ and $l \in N \setminus S$: $q^*_{ki} = q^*_{li}$ ($q^*_{kj} = q^*_{lj}$). In every signatory market $k \in S$: $q^*_{ki} \leq q^*_{kj}$; in every non-signatory market $l \in S \setminus N$: $q^*_{li} < q^*_{lj}$. Regarding total production: $Q^*_{.i} < Q^*_{.j}$. Regarding total consumption: in all markets $k \in S$ and $l \in S \setminus N : Q^*_{k.} = Q^*_{l.}$.

BCA-regime: signatory firm $i's$ outputs for signatory markets $k \in S$ compared to their outputs for non-signatory markets $l \in S \setminus N$: $q^*_{ki} > q^*_{li}$; except for $\gamma = 0$ in which case $q^*_{ki} = q^*_{li}$. Non-signatory firm $j's$ outputs for non-signatory markets $l \in S \setminus N$ compared to their outputs for signatory markets $k \in S$: $q^*_{lj} > q^*_{kj}$. In every signatory market $k \in S$: $q^*_{ki} = q^*_{kj}$; in every non-signatory market $l \in S \setminus N$: $q^*_{li} < q^*_{lj}$. Regarding total production: $Q^*_{.i} < Q^*_{.j}$. Regarding total consumption in every signatory market $k \in S$ and in every non-signatory market $l \in S \setminus N$: $Q^*_{k.} < Q^*_{l.}$.

Appendix 2

In stage 2, quantities as a function of taxes as known from stage three are inserted into governments' welfare functions and the welfare functions are differentiated with respect to own taxes. Importantly, different from the quantities displayed in Appendix 1, which already use the information of symmetry, quantities must be expressed as a function of the full tax vector, and the symmetry assumption can only be invoked after derivatives have been taken. The simultaneous solution of the m first order conditions of signatory governments and $n-m$ first order conditions of non-signatory governments delivers equilibrium taxes, which are a function of all parameters of the model, $t_i^*(m,a,c,\delta)$ and $t_j^*(m,a,c,\delta)$.

Inserting equilibrium taxes in equilibrium quantities, as displayed in Appendix 1, reveals that we need to impose constraints on the parameters such that equilibrium quantities are positive. Such non-negativity constraints essentially boil down to the condition that the demand parameter a is larger than marginal production cost c plus a multiple of marginal damages δ. In other words, these non-negativity constraints represent a lower threshold $\underline{a}$, such that if $a \geq \underline{a}$ holds, all quantities are positive for all m, $1 \leq m \leq n$.

Typically, the most restrictive condition applies in the grand coalition ($m=n$) with the highest overall tax level. For instance, under the No BCA-regime, the non-negativity constraint is $a \geq \underline{a} = c + m\delta$ for $\gamma = 0$ and $a \geq \underline{a} = c + n\delta\Psi_A/m\Psi_B$ with $\Psi_A = n^2(m-1) - n(m-1)^2 - m(m-2)$, $\Psi_B = n^2 - (m-1)(n+1)$, $\boldsymbol{\Psi_A} > 0$ and $\Psi_B > 0$ for $\gamma = 1$. For $\gamma = 0$ it is immediately evident that the condition is most restrictive if $m = n$. However for $\gamma = 1$, this is also the case because one can show that $n\Psi_A/m\Psi_B$ increases in m. Hence, by setting $m = n$, the non-negativity constraint is given by $a > c + n\delta$ for both values of γ. An interpretation of this condition is that global marginal damages $n\delta$ cannot be too high compared to the term $a-c$, which can be interpreted as the market size corrected for production costs. The term $a-c$ can also be viewed as the benefit from production and consumption.

For the BCA-regime, similar non-negativity constraints can be derived, even though they look much more complicated but in essence they also require that global marginal damages cannot be too high compared to the "corrected" market size.

Another issue, which is important for our analysis, is that we want to consider situations for which $t_i^* \geq t_j^*$ holds. In our model (like in many other strategic trade models; see, e.g., Benchekroun and van Long, 1998), this is not automatically guaranteed. The reason is simple: in this game, we have two market imperfections. On the one hand, there is Cournot-competition in international trade. This implies that signatories, which internalize externalities among their members, have an incentive to subsidize their consumers. However, they also have an incentive to tax their producers in order to enforce a cartel solution, i.e., stabilizing the market price by reducing output and output is reduced through taxes. Different from models with only two countries and no coalition formation, in our model signatory governments have an incentive to tax their producers. The reason is simple. In our model, taxes, which producers have to pay, are welfare neutral as they constitute revenues for the government. The gross profits of firms (i.e., profits excluding tax bills) are maximized in a monopoly. As firms compete non-cooperatively in each market as long as $\gamma > 0$, signatory governments aiming at maximizing gross profits of their firms have an incentive to collude. Trading off firms' gross profits for consumer surplus leads to a subsidy if the market size is large compared to production cost, represented by the parameters a and c in our model. On the other hand, there is global pollution, which calls for taxes in order to reduce damages. The importance of damages in the welfare function is represented by the parameter δ in our model. Hence, as known from many strategic trade models, even if the grand coalition forms, taxes may not be set at the Pigouvian level, i.e., not being equal to the sum of marginal damages $n\delta$. They may be set lower or higher, depending on the importance consumers and producers receive in governments' welfare function. Thus, if the value of the parameter a is high compared to the value of the parameters c and δ, signatory governments may choose a lower tax than non-signatory governments and taxes may even be negative, i.e., governments subsidize their firms. For instance, under the No BCA-regime and $\gamma = 0$, $t_i^*(m) \geq t_j^*(m) > 0$ if $a \leq c + 2n\delta$ but $t_i^*(m) < t_j^*(m) < 0$ if $a > c + 2n\delta$, both inequalities are compatible with the non-negativity constraint derived above, i.e., $a \geq c + n\delta$. For $\gamma = 1$, one can show that $t_i^*(m) > t_j^*(m)$ and $t_j^*(m) < 0$ always hold whereas $t_i^*(m)$ can be positive or negative without violating the non-negativity constraint.

The outcome of all this is that in the context of our analysis it makes sense to impose an upper bound threshold on a, denoted by $\overline{a}$ such that if $a \leq \overline{a}$, $t_i^*(m) \geq t_j^*(m)$ is true. In other words, we assume damages to be significant enough for governments such that signatories choose higher taxes than non-signatories, which seems to be the basic motivation to consider border tax adjustments for environmental reasons. Thus, even under the No BCA-regime we assume damages to be sufficiently strong such that signatories choose higher taxes than non-signatories, $t_i^*(m) \geq t_j^*(m)$. For the No BCA-regime and $\gamma = 0$, this implies for instance $\overline{a} = c + 2n\delta$. In the same spirit, also under the BCA-regime, we can derive upper bound thresholds $\overline{a}$ such that if $a \leq \overline{a}$, $t_i^*(m) \geq t_j^*(m)$ is always true.

Clearly, when comparing the two regimes, we need to assume that parameter a satisfies both non-negativity constraints, thus the joint lower bound $\underline{a}$ is the maximum of the two lower bounds. Also, regarding the upper bound $\overline{a}$, the joint upper bound is the minimum of the two upper bounds. Generally, both bounds depend on γ, δ and c. However, as we assume $c = 0$ in the simulations, we just write $\underline{a}(\gamma, \delta)$ and $\overline{a}(\gamma, \delta)$ below.

Simulations have been conducted with the mathematical software program Maple, and all files and detailed numerical results are available upon request from the authors. We consider three values of the TFV parameter: no TFV with $\gamma = 1$, partial TFV with $\gamma = 0.5$, and full TFV with $\gamma = 0$. We also consider three values for the damage parameter $\delta = \{10, 50, 100\}$, assume $n = 10$ countries, and normalize the cost parameter by setting $c = 0$. We recall that we assume $\phi = 1$ for the adjustment parameter. For these parameter values, the lower ($\underline{a}$) and upper ($\overline{a}$) bounds of parameter a are listed in the table below.

	$\delta = 10$		$\delta = 50$		$\delta = 100$	
	$\underline{a}$	$\overline{a}$	$\underline{a}$	$\overline{a}$	$\underline{a}$	$\overline{a}$
$\gamma = 0$	101	104	501	520	1001	1040
$\gamma = 0.5$	120	1000	590	5300	1250	10500
$\gamma = 1$	250	5000	1350	25000	2500	50000

Each parameter space of parameter a, i.e., $\underline{a}(\gamma, \delta) \leq a(\gamma, \delta) \leq \overline{a}(\gamma, \delta)$, is divided into 6 a-values with equidistant intervals such that $\triangle(\gamma, \delta) = \frac{\overline{a}(\gamma,\delta)-\underline{a}(\gamma,\delta)}{5}$, $a_1(\gamma, \delta) = \underline{a}(\gamma, \delta)$, $a_2(\gamma, \delta) = \underline{a}(\gamma, \delta) + \triangle(\gamma, \delta)$, ..., $a_6(\gamma, \delta) = \underline{a}(\gamma, \delta) + 5\triangle(\gamma, \delta) = \overline{a}(\gamma, \delta)$. Thus, for instance, for $\gamma = 0$ and $\delta = 10$, we have $a_1(\gamma, \delta) = 101$, $a_2(\gamma, \delta) = 101.6$, ..., $a_6(\gamma, \delta) = 104$. In general, for a given γ, moving from $\delta = 10$ to $\delta = 50$, all bounds of parameter a are inflated by a factor of 5 and by moving to $\delta = 100$, they are inflated by a factor of 10. Thus, the different values of the damage parameter δ can be considered as a sensitivity analysis for which all qualitative results on which we report turn out to be robust. (Therefore, absolute values of the parameters a, c and δ do not matter but their ratio.)

Finally, in all simulations, we test for second derivatives. It turns out that welfare functions are strictly concave in own taxes, $\partial^2 W_{i \in S}(t_i)/\partial t_i^2 < 0$, $\partial^2 W_{j \notin S}(t_j)/\partial t_j^2 < 0$, and that the Hessian matrix is semi-definite, guaranteeing a unique stable equilibrium tax vector.

All results reported in Section 5 are detailed in our working paper in Appendix 3. Further details of all Maple files used for the simulations are available upon request from the authors.

Tables

Table 1: Stable Coalitions and Global Welfare under Open Membership

	No BCA-Regime		BCA-Regime	
γ	m^*	$CGI(m^*)$	m^*	$CGI(m^*)$
1	1	0	10	100
0.5	2	1	10	100
0	3	7.3	6-9	87.2-99.9

Holds for a simultaneous and sequential coalition formation process; m^* denotes the size of stable coalitions under open membership and $CGI(m^*)$ is the associated closing the gap index as defined in Eq. (11) in the text.

Table 2: Stable Coalitions and Global Welfare under Exclusive Membership and Sequential Coalition Formation

	Parameter a	No BCA-Regime		BCA-Regime	
γ		m^{**}	$CGI(m^{**})$	m^{**}	$CGI(m^{**})$
1	$a_1(\delta,\gamma) / a_2(\delta,\gamma) / a_3(\delta,\gamma)$	1	0	9/6/6	98.7/93.5/44.1
	$a_4(\delta,\gamma) / a_5(\delta,\gamma) / a_6(\delta,\gamma)$	1	0	5/5/5	<0/<0/<0
0.5	$a_1(\delta,\gamma) / a_2(\delta,\gamma) / a_3(\delta,\gamma)$	2	1	9/8/6	97.3/95.5/78.9
	$a_4(\delta,\gamma) / a_5(\delta,\gamma) / a_6(\delta,\gamma)$	2	1	6/5/5	<0/<0/<0
0	$a_1(\delta,\gamma) / a_2(\delta,\gamma) / a_3(\delta,\gamma)$	3	7.3	6/8/8	87.9/97.7/98.0
	$a_4(\delta,\gamma) / a_5(\delta,\gamma) / a_6(\delta,\gamma)$	3	7.3	9/9/9	99.8/99.9/99.9

m^{**} denotes the size of stable coalitions and $CGI(m^{**})$ the associated closing the gap index as defined in Eq. (11) in the text. $a_k(\delta,\gamma)$ are the six values as defined in Appendix 2. $CGI(m^{**})$: <0 means a negative CGI, CGI-values for the No-BCA regime are the same for all parameter values, CGI for BCA regime assumes the particular value $\delta = 10$ with very similar values for $\delta = 50$ and $\delta = 100$.

Figure 1a: Welfare of Signatories and BCAs

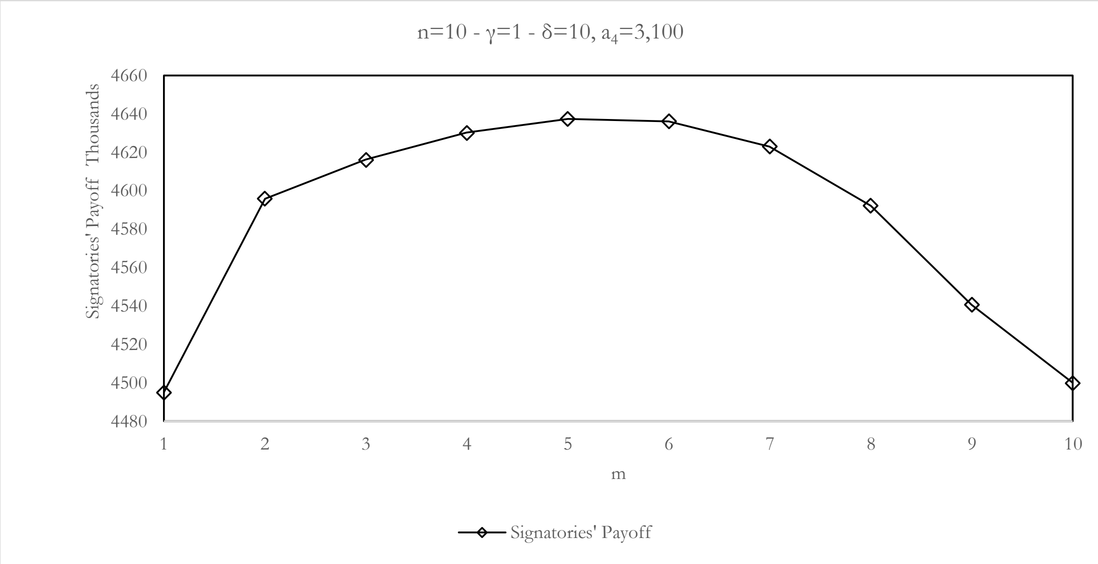
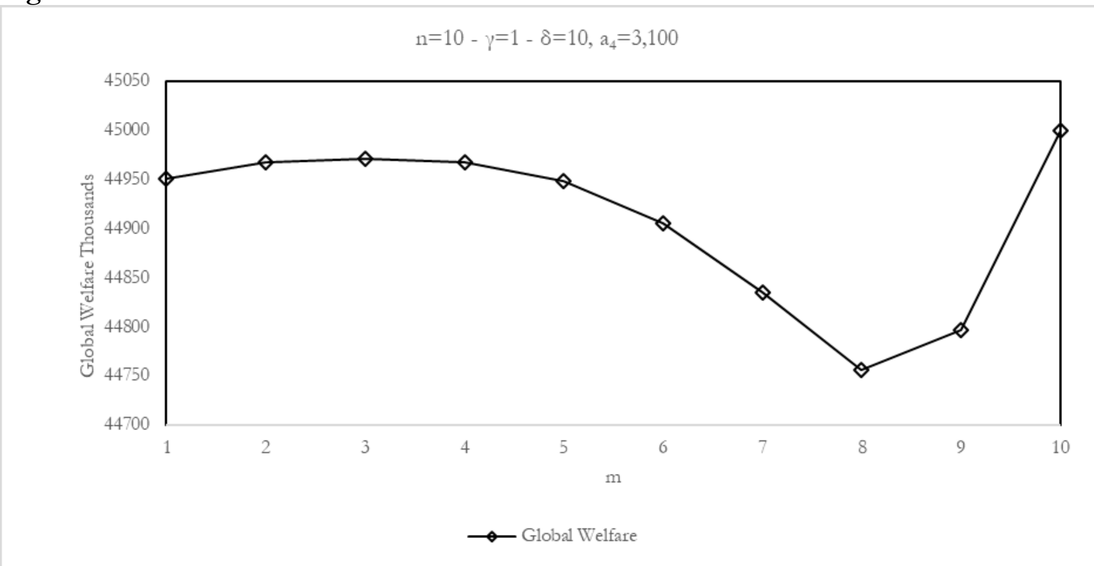


Figure 1b: Global Welfare and BCAs



Graz Economics Papers

For full list see:
http://ideas.repec.org/s/grz/wpaper.html
Address: Department of Economics, University of Graz,
Universitätsstraße 15/F4, A-8010 Graz

04–2020	**Alaa Al Khourdajie and Michael Finus**: Measures to Enhance the Effectiveness of International Climate Agreements: The Case of Border Carbon Adjustments
03–2020	**Katja Kalkschmied**: Rebundling Institutions
02–2020	**Miriam Steurer and Caroline Bayr**: Measuring Urban Sprawl using Land Use Data
01–2020	**Thomas Aronsson, Sugata Ghosh and Ronald Wendner**: Positional Preferences and Efficiency in a Dynamic Economy
14–2019	**Nicole Palan, Nadia Simoes, and Nuno Crespo**: Measuring Fifty Years of Trade Globalization
13–2019	**Alejandro Caparrós and Michael Finus**: Public Good Agreements under the Weakest-link Technology
12–2019	**Michael Finus, Raoul Schneider and Pedro Pintassilgo**: The Role of Social and Technical Excludability for the Success of Impure Public Good and Common Pool Agreements: The Case of International Fisheries
11–2019	**Thomas Aronsson, Olof Johansson-Stenman and Ronald Wendner**: Charity as Income Redistribution: A Model with Optimal Taxation, Status, and Social Stigma
10–2019	**Yuval Heller and Christoph Kuzmics**: Renegotiation and Coordination with Private Values
09–2019	**Philipp Külpmann and Christoph Kuzmics**: On the Predictive Power of Theories of One-Shot Play
08–2019	**Enno Mammen, Jens Perch Nielsen, Michael Scholz and Stefan Sperlich**: Conditional variance forecasts for long-term stock returns

07–2019 **Christoph Kuzmics, Brian W. Rogers and Xiannong Zhang**: Is Ellsberg behavior evidence of ambiguity aversion?

06–2019 **Ioannis Kyriakou, Parastoo Mousavi, Jens Perch Nielsen and Michael Scholz**: Machine Learning for Forecasting Excess Stock Returns The Five-Year-View

05–2019 **Robert J. Hill, Miriam Steurer and Sofie R. Waltl**: Owner-Occupied Housing, Inflation, and Monetary Policy

04–2019 **Thomas Aronsson, Olof Johansson-Stenman and Ronald Wendner**: Charity, Status, and Optimal Taxation: Welfarist and Paternalist Approaches

03–2019 **Michael Greinecker and Christoph Kuzmics**: Limit Orders under Knightian Uncertainty

02–2019 **Miriam Steurer and Robert J. Hill**: Metrics for Measuring the Performance of Machine Learning Prediction Models: An Application to the Housing Market

01–2019 **Katja Kalkschmied and Jörn Kleinert**: (Mis)Matches of Institutions: The EU and Varieties of Capitalism

21–2018 **Nathalie Mathieu-Bolh and Ronald Wendner**: We Are What We Eat: Obesity, Income, and Social Comparisons

20–2018 **Nina Knittel, Martin W. Jury, Birgit Bednar-Friedl, Gabriel Bachner and Andrea Steiner**: The implications of climate change on Germanys foreign trade: A global analysis of heat-related labour productivity losses

19–2018 **Yadira Mori-Clement, Stefan Nabernegg and Birgit Bednar-Friedl**: Can preferential trade agreements enhance renewable electricity generation in emerging economies? A model-based policy analysis for Brazil and the European Union

18–2018 **Stefan Borsky and Katja Kalkschmied**: Corruption in Space: A closer look at the world's subnations

17–2018 **Gabriel Bachner, Birgit Bednar-Friedl and Nina Knittel**: How public adaptation to climate change affects the government budget: A model-based analysis for Austria in 2050

16–2018 **Michael Günther, Christoph Kuzmics and Antoine Salomon**: A Note on Renegotiation in Repeated Games [Games Econ. Behav. 1 (1989) 327360]

15–2018 **Meng-Wei Chen, Yu Chen, Zhen-Hua Wu and Ningru Zhao**: Government Intervention, Innovation, and Entrepreneurship

14–2018 **Yu Chen, Shaobin Shi and Yugang Tang**: Valuing the Urban Hukou in China: Evidence from a Regression Discontinuity Design in Housing Price

13–2018 **Stefan Borsky and Christian Unterberger**: Bad Weather and Flight Delays: The Impact of Sudden and Slow Onset Weather Events

12–2018 **David Rietzke and Yu Chen**: Push or Pull? Performance-Pay, Incentives, and Information

11–2018 **Xi Chen, Yu Chen and Xuhu Wan**: Delegated Project Search

10–2018 **Stefan Nabernegg, Birgit Bednar-Friedl, Pablo Muñoz, Michaela Titz and Johanna Vogel**: National policies for global emission reductions: Effectiveness of carbon emission reductions in international supply chains

09–2018 **Jonas Dovern and Hans Manner**: Order Invariant Tests for Proper Calibration of Multivariate Density Forecasts

08–2018 **Ioannis Kyriakou, Parastoo Mousavi, Jens Perch Nielsen and Michael Scholz**: Choice of Benchmark When Forecasting Long-term Stock Returns

07–2018 **Joern Kleinert**: Globalization Effects on the Distribution of Income

06–2018 **Nian Yang, Jun Yang and Yu Chen**: Contracting in a Continuous-Time Model with Three-Sided Moral Hazard and Cost Synergies

05–2018 **Christoph Kuzmics and Daniel Rodenburger**: A case of evolutionary stable attainable equilibrium in the lab

04–2018 **Robert J. Hill, Alicia N. Rambaldi, and Michael Scholz**: Higher Frequency Hedonic Property Price Indices: A State Space Approach

03–2018 **Reza Hajargasht, Robert J. Hill, D. S. Prasada Rao, and Sriram Shankar**: Spatial Chaining in International Comparisons of Prices and Real Incomes

02–2018 **Christoph Zwick**: On the origin of current account deficits in the Euro area periphery: A DSGE perspective

01–2018 **Michael Greinecker and Christopher Kah**: Pairwise stable matching in large economies