

Strategic Environmental Policy, International Trade and Self-enforcing Agreements: The Role of Consumers’ Taste for Variety

Michael Finus¹ and Alaa Al Khourdajie^{2*}

¹*Department of Economics, University of Bath, Bath, BA2 7AY, UK;
m.finus@bath.ac.uk*

²*School of Economics, University of Edinburgh, Edinburgh, EH8 9JT,
UK; alaa.alkhourdajie@ed.ac.uk*

ABSTRACT

We study the coordination of environmental policy within an agreement in the context of international trade. In an n-country intra-industry trade model, firms produce a horizontally differentiated good and consumers have a taste for variety. Governments choose strategically an emission tax and their membership in an international agreement. We show that only a strong taste for variety reduces the competition among governments sufficiently enough to allow for some form of policy coordination, though full cooperation will never be obtained.

Keywords: Strategic environmental policy, international trade, self-enforcing international agreements, horizontal product differentiation, taste for variety

JEL Codes: C72, F18, Q58

*We are grateful for the feedback of the attendees of the first AERNA Workshop (Spanish–Portuguese Association of Resource and Environmental Economics) on Game Theory and the Environment, Department of Economic Analysis, Universidad

ISSN 1944-012X; DOI 10.1561/102.00000086

© 2018 M. Finus & A. Al Khourdajie

Introduction

Reaching a meaningful international agreement on climate change has proved difficult over the last three decades. The Kyoto Protocol, signed in 1998, could not even stop the trend of a continuous increase of greenhouse gas emissions world-wide observed since the last century. In the most recent round of climate change negotiations in Paris in December 2015, even though many countries around the world signed an agreement, it is only based on voluntary pledges of governments, without any enforcement mechanism in case of non-compliance. Moreover, even if all governments would deliver on their pledges, global temperature is expected to increase by 2.7–3°C (UNFCCC, 2015), much above the widely accepted and recommended target of limiting the temperature increase by 2100 to 2°C compared to per-industrial levels.

Scholars in the game-theoretic literature on the formation of self-enforcing international environmental agreements (IEAs) attribute the difficulty in reaching an effective climate change agreement to strong free-riding incentives. These incentives emerge because any non-signatory can enjoy the environmental benefits from reduced emissions without incurring any cost. In the absence of supranational authority that could enforce cooperation on climate change, self-enforcing agreements achieve relatively little. A central finding of this literature is that either participation in an agreement is small or if it is large, then the difference between cooperative and non-cooperative behavior is small, that is, cooperation does not really matter. Barrett (1994a) called this the paradox of cooperation.¹ For a recent survey of the literature, including a collection of the most influential papers over the last two decades, see Finus and Caparrós (2015).

Autonoma de Madrid, Spain, September 2015. We are also grateful for the feedback of the attendees of the 2016 Annual Conference of the European Association of Environmental and Resource Economists in Zurich, Switzerland. We also would like to acknowledge helpful comments by Javier Rivas and Chuck Mason as well as by two anonymous referees who provided excellent comments on a previous version of the paper. Needless to say, the authors assume the responsibility of any remaining errors and omissions.

¹More optimistic results have been derived for instance for modest emission reductions (Barrett, 2002; Finus and Maus, 2008) and if countries cooperate on R&D instead of mitigation (Barrett, 2006; El-Sayed and Rubio, 2014; Hoel and Zeeuw, 2010).

Another body of literature explaining the slow progress in addressing transboundary pollution problems, in particular climate change, points at the fear of governments to loose competitiveness in international trade if they pursue a stricter environmental policy than other governments (Copeland and Taylor, 2005). Based on an extension of the simple strategic trade policy model of Brander and Spencer (1985), strategic environmental policy has been analyzed for instance by Barrett (1994b), Conrad (1993), and Kennedy (1994). Under Cournot-competition,² Brander and Spencer have shown that governments have an incentive to subsidize production of own firms in order to increase their rent capture. For environmental policy this means that emission taxes are set below marginal damages (Barrett, 1994b). This result has been modified in several directions by considering additional components in governments' welfare function. Adding consumers to such a model lowers equilibrium environmental taxes even further because consumers call for larger quantities in a model of imperfect competition (Kennedy, 1994). Similarly, departing from the assumption of a local pollutant and considering transboundary pollution provides further incentives to lower environmental taxes. This is because governments have an incentive to externalize some environmental damages, understanding that domestic production is substituted by foreign production if heavily taxed, which may even increase environmental damages if foreign production is more dirty than domestic production (Conrad, 1993). Taken together, strategic trade models offer a rich setting to explain why environmental policies may be distorted, which is also evident from Ulph (1996a,b), considering also the incentive of firms to strategically invest in R&D and by allowing governments to use different environmental policy instruments.

Essentially for a long time, both strands of literature have not been integrated. That is, the IEA literature did not explicitly consider trade and the strategic environmental policy and the trade literature did not allow for the formation of agreements, namely, it did not consider the possibility that governments coordinate their policies. Only recently, Eichner and Pethig produced a series of papers considering both aspects, Eichner and Pethig (2012, 2013, 2014, 2015), though their trade model

²Barrett (1994b) has shown that, probably not surprisingly, many of the results reverse if Bertrand-competition is considered.

is very different from those mentioned earlier, and hence their results are difficult to relate to this literature. Eichner and Pethig extend the basic model of international environmental agreements without trade (Barrett, 1994a) by introducing production, consumption, and international trade. The authors assume composite consumer goods and fossil fuel goods that are produced and consumed in each country and traded in international markets. Furthermore, they assume that consumers face carbon caps imposed on their consumption of fossil fuel goods. Eichner and Pethig (2012) find that stable coalitions are small and hence ineffective in reducing emissions. In Eichner and Pethig (2013), they consider that the coalition acts like a Stackelberg leader. Hence, in line with Barrett (1994a), they find that stable coalitions can be larger. However, due to emission leakage, equilibrium emissions are only slightly lower than in the business as usual scenario, rendering the environmental benefits of these coalitions negligible. Eichner and Pethig (2015) extend their previous analysis by considering carbon taxes instead of emission caps. They find that the grand coalition can be stable. However, this is only the case if the damages from pollution are low, for medium to high level of damages, the grand coalition is no longer stable. Thus, they essentially confirm the paradox of cooperation. The interesting aspect is that the policy instrument matters for stability. Finally, in Eichner and Pethig (2014) they consider asymmetric countries and a carbon tax. They consider two types of asymmetry, in terms of pollution damages and fossil fuel demand. They find that in the case of asymmetric damages, the grand coalition is never stable, unless the asymmetry as well as the damage levels are very low. In the case of a very high degree of fuel demand asymmetry, the grand coalition can be stable, but only if fuels are very scarce worldwide.

In contrast, we offer an IEA-model which is very much in the spirit of the strategic environmental policy and trade literature. Our model considers governments which care for the profits of their firms, the utility of their consumers and environmental damages, which are the result of a global pollutant. They choose strategically an emission tax and their membership in an international agreement. We allow for horizontal product differentiation where consumers' taste for variety is captured. Thus, our paper benefits from contributions by Yi (1996, 2000) and Loke and Winters (2012) who look at international trade,

trade agreements and taste for variety, but who ignore environmental damages and their effect on governments' strategic behavior.

Adding agreement formation to a strategic trade and environmental model changes some of the incentives mentioned earlier. First of all and foremost, members to an agreement can internalize, at least partially, some of the externalities, affecting firms, consumers, and environmental damages. To which extent this undertaking is successful depends on the strategic interaction of environmental policy between signatory and non-signatory countries and the size of stable agreements. Secondly, the rent capture argument for low taxes is reversed in a game with agreement formation. Given that taxes are welfare neutral in such type of models (tax bills by firms are equal to tax revenues of governments), in a market with oligopolistic competition, signatory governments have an incentive to enforce a cartel solution for their firms, that is, lowering output in order to increase the market price via an increase in taxes. Internalizing environmental damages among members to an agreement also calls for even higher taxes. In contrast, if members to an agreement care for their consumers sufficiently enough will they have an incentive to lower their taxes compared to non-members. Overall, our results confirm the predominantly pessimistic conclusion which has emerged from the IEA-literature. Agreements are small at best and may not exist at all due to strong free-rider incentives. If consumers have a low taste for variety, that is, domestic and foreign varieties are viewed as good substitutes by consumers, agreement formation fails. Only with a sufficiently high taste for variety, strategic interaction of governments is sufficiently reduced such that small agreements are stable.

In what follows, the following section presents our model. The next section develops our results, including an in-depth analysis of the driving forces of coalition formation and the strategic interaction between signatories and non-signatories to an agreement. In the last section, we summarize our main results and conclude.

Model

Payoff Function

Consider an intra-industry trade model with n *ex ante* symmetric countries with a representative firm and consumer in each country.

We denote the set of countries by N . Firms produce a horizontally differentiated good, that is, the same good but in different varieties where each firm produces one variety. This good causes environmental damages; the production uses emissions as an input, for example in the form of energy by which greenhouse gases are released. Firms compete in a 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.³ Transport costs are assumed away as usual.

The welfare of country i is given by:

$$W_i = PS_i + TR_i + CS_i - D_i \quad (1)$$

where PS_i is country i 's producer surplus, TR_i is the tax revenue from the emission tax imposed by the government i on its domestic firm, CS_i represents country i 's consumer surplus, and D_i is the pollution damage faced by country i .

Consumers are identical and their preferences are represented by a quasi-linear utility function over two goods (see Equation (2)). The first good is the horizontally differentiated and traded good. The second good is a numeraire good, representing the composition of all other goods. Utility is linear in the numeraire good and quadratic in the differentiated good.

We assume that consumers have a taste for variety (Dixit and Stiglitz, 1977). That is, their utility depends not only on the total quantity consumed but also on the composition of quantities of the differentiated good (Yi, 1996, 2000). The taste for variety (abbreviated TFV hereafter) is captured by parameter $\gamma \in [0, 1]$. High values of γ imply a low taste for variety and for $\gamma = 1$ varieties are perfect substitutes. In contrast, low values of γ represent a high preference for a diverse and balanced consumption bundle and for $\gamma = 0$ varieties cannot be substituted at all.⁴

³See Appleyard and Field (2014) as well as Helpman and Krugman (1985) for further background.

⁴An 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 bias towards the domestically produced variety (Di Comite *et al.*, 2014).

More specifically, let the representative consumer's utility in country i be given by u_i :

$$u_i(q_i; M_i) = v_i(q_i) + M_i = aQ_i - \frac{\gamma}{2}Q_i^2 - \frac{1-\gamma}{2} \sum_{k \in N} q_{ik}^2 + M_i \quad (2)$$

where v_i represents the utility from consuming the horizontally differentiated and traded good and M_i represents the utility from consuming the numeraire good; $q_i = (q_{i1}, \dots, q_{in})$ is a vector of varieties consumed by consumers in country i , with q_{ik} representing country i 's consumption of country k 's variety⁵; a is a positive demand parameter and $Q_i = \sum_{k \in N} q_{ik}$ is country i 's total consumption of all varieties, supplied by all countries k .

In this paper, in most parts, we will focus our analysis on two extreme TFV scenarios: the "no TFV" scenario with $\gamma = 1$ and the "maximum TFV" scenario with $\gamma = 0$. This is done for analytic tractability and can be justified because the driving forces identified below for these two extreme assumptions would also be present for intermediate values, though with different degrees.

From Equation (2), country i 's inverse demand function for country k 's variety follows from:

$$p_{ik} = \frac{\partial u_i}{\partial q_{ik}} \iff p_{ik} = a - (1-\gamma)q_{ik} - \gamma Q_i \iff p_{ik} = a - q_{ik} - \gamma \sum_{l \in N, l \neq k} q_{il} \quad (3)$$

where p_{ik} represents the price faced by consumers in country i consuming the variety of country k and $\sum_{l \in N, l \neq k} q_{il}$ is the sum of all consumed varieties produced by all firms except firm k in country i .

From Equation (2) and (3), the consumer surplus in country i is given by :

$$CS_i = aQ_i - \frac{\gamma}{2}Q_i^2 - \frac{1-\gamma}{2} \sum_{k \in N} q_{ik}^2 - \sum_{k \in N} q_{ik}p_{ik} \quad (4)$$

where the last term in Equation (4) represents consumers' spending.

⁵Throughout the paper the first subscript indicates the market in which the variety is consumed and the second subscript indicates the market in which it is produced.

The producer surplus of representative firm i (in country i) is the sum of its profits in each market:

$$PS_i = \sum_{k \in N} \pi_{ki} = \sum_{k \in N} q_{ki}(p_{ki} - c - t_i) \quad (5)$$

where π_{ki} is firm i 's profit in market k from selling quantity q_{ki} at price p_{ki} where c is the constant marginal cost and t_i is the emission tax imposed by country i 's government on its firm's output, which assumes that emissions are linked to quantities by a constant emission-output coefficient. Without loss of generality, we set this coefficient to 1. This implies that in this simple model, the tax is de facto an output tax. Moreover, the tax is an efficient instrument to tackle all externalities present in this model, as all welfare components depend directly on outputs. Also the environmental externality is efficiently tackled by this tax because the only option to address environmental externalities is to reduce output. This would be different in richer models (e.g., Conrad, 1993) in which firms can reduce emissions by investing in a different production/abatement technology. Given imperfect competition, the tax is typically not set equal to marginal damages.

The tax revenue, TR_i , is given by:

$$TR_i = t_i \sum_{k \in N} q_{ki} \quad (6)$$

and damages from global pollution faced by country i are given by:

$$D_i = \delta \sum_{i \in N} Q_i. \quad (7)$$

where δ is a damage parameter, $\sum_{i \in N} Q_i$ is total consumption in every country i and hence total emissions (due to our assumption of a constant emission output coefficient of 1). That is, emissions constitute a pure public bad: damages depend on total emissions.

Coalition Formation Game

We assume a three-stage coalition formation game, which unfolds as follows.

Stage 1, Choice of Membership: all countries decide simultaneously whether to join coalition S with m the cardinality of S . Countries

which do not join S act as singletons. A typical signatory will be denoted by i and a non-signatory by j .

Following d'Aspremont *et al.* (1983), a coalition is called stable if it is internally and externally stable. Internal stability means that no signatory has an incentive to leave coalition S , whereas external stability means that no non-signatory has an incentive to join coalition S . We assume for simplicity that in the case of indifference a non-signatory joins coalition S .

Internal stability:

$$W_i(S) - W_i(S \setminus \{i\}) \geq 0 \quad \forall i \in S \quad (8)$$

External stability:

$$W_j(S) - W_j(S \cup \{j\}) > 0 \quad \forall j \in N \setminus S. \quad (9)$$

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

- Signatories choose their joint emission tax t_i (implemented uniformly in all signatory countries) in order to maximize the joint welfare of coalition S : $\max_{t_i} \sum_{i \in S} W_i$
- Non-signatories choose their individual tax t_j in order to maximize their individual welfare: $\max_{t_j} W_j$.

That is, the group of signatories with m members and all $n - m$ non-signatories play a Nash equilibrium among each other.

Stage 3, Choice of Output: all firms choose simultaneously and non-cooperatively their segmented market outputs by maximizing profits:

$$\max_{q_{1i}, \dots, q_{ni}} PS_i.$$

That is, firms play a Nash equilibrium among each other. The game is solved by backwards induction.

Results

Third Stage

In this section, we derive results for the third stage. The profit of firm i in market k is given by $\pi_{ki} = q_{ki}(p_{ki} - c - t_i)$. Substituting the inverse

demand function from Equation (3), we derive the following first-order condition:

$$\begin{aligned} \frac{\partial \pi_{ki}}{\partial q_{ki}} &= a - c - t_i - (2 - \gamma)q_{ki} - \gamma Q_k = 0 \\ \iff a - c - t_i - 2q_{ki} - \gamma \sum_{l \in N, l \neq i} q_{kl} &= 0 \end{aligned} \quad (10)$$

where Q_k is the total quantity consumed in market k and $\sum_{l \in N, l \neq i} q_{kl}$ is the sum of all consumed varieties by consumers in market k from all firms except from firm i . It is easy to see that reaction functions ($q_{ki} = r_i(\sum_{l \in N, l \neq i} q_{kl})$) have a slope of $-\gamma/2$. Hence, the equilibrium is unique; the absolute value of the slope of the reaction function increases with the taste of variety parameter γ and as γ approaches zero, the strategic interaction among firms vanishes. Moreover, it is apparent that a necessary condition for positive quantities is $a > c$. Below, we will further develop this non-negativity condition in order to ensure interior solutions.

Solving the n first-order conditions in market k simultaneously, gives⁶

$$q_{ki} = \frac{(a - c)(2 - \gamma) - t_i(\gamma(n - 2) + 2) + \gamma \sum_{l \in N, l \neq i} t_l}{(\gamma(n - 1) + 2)(2 - \gamma)}. \quad (11)$$

Since the tax is imposed on production, the equilibrium quantity of firm i 's variety is the same in all markets k . It is evident that quantities decrease in own taxes and increase in foreign taxes. If we already account for the fact of a symmetric tax equilibrium in stage 2 with all signatories choosing the same tax rate t_i and all non-signatories choose the same tax rate t_j (and typically $t_i \neq t_j$), then we have for a signatory's firm

$$q_{i \in S}^* = \frac{(a - c)(2 - \gamma) - t_i(\gamma(n - m - 1) + 2) + t_j(\gamma(n - m))}{(\gamma(n - 1) + 2)(2 - \gamma)} \quad (12)$$

⁶In market k we have one first-order condition for firm i : (a) $a - c - t_i - (2 - \gamma)q_{ki} - \gamma Q_k = 0$ and $n - 1$ first-order conditions for all other firms: (b) $a - c - t_l - (2 - \gamma)q_{kl} - \gamma Q_k = 0$. Summing over all n firms, we derive aggregate supply: $Q_k^* = \frac{n(a - c) - t_i - \gamma \sum_{l \neq i} t_l}{\gamma(n - 1) + 2}$. Substituting back in (a), and solving for q_{ki} gives for instance (11).

and for a non-signatory's firm

$$q_{j \notin S}^* = \frac{(a - c)(2 - \gamma) + \gamma m t_i - t_j(\gamma(m - 1) + 2)}{(\gamma(n - 1) + 2)(2 - \gamma)} \quad (13)$$

with the total equilibrium consumption in market k , $Q_{k.}^*$ given by:

$$Q_{k.}^* = \frac{n(a - c) - t_j(n - m) - m t_i}{\gamma(n - 1) + 2} \quad (14)$$

This leads to the following conclusions.

Proposition 1 (The Effects of Taxes on Equilibrium Quantities). *Consider the third stage and a market k . Suppose a coalition S has formed in the first stage and all players have chosen their taxes in stage 2, with all m signatories choosing t_i and all $n - m$ non-signatories choosing t_j . The quantity of firm i 's (j 's) variety in a signatory country (non-signatory country) decreases with the level of signatories' (non-signatories') taxes, $\frac{\partial q_{i.}}{\partial t_i} < 0$ ($\frac{\partial q_{j.}}{\partial t_j} < 0$) and increases with the level of non-signatories' (signatories') taxes, $\frac{\partial q_{i.}}{\partial t_j} > 0$ ($\frac{\partial q_{j.}}{\partial t_i} > 0$), except for $\gamma = 0$ in which case $\frac{\partial q_{i.}}{\partial t_j} = 0$ ($\frac{\partial q_{j.}}{\partial t_i} = 0$). The total quantity in market k decreases in signatories' and non-signatories taxes, $\frac{\partial Q_{k.}^*}{\partial t_i} < 0$ and $\frac{\partial Q_{k.}^*}{\partial t_j} < 0$ irrespective of γ .*

Proof. Follows directly from Equations (12) to (14). $\square$

Thus, quantities produced by a firm for a particular market are negatively affected by own taxes and positively affected by foreign taxes. Given that a firm produces the same quantities for all markets, also the same holds for total production of a firm. Only for the maximum TFFV, $\gamma = 0$, will a firm's output not be affected by the tax of a foreign government imposed on a foreign firm. Then, essentially, firms act in each segmented market like a monopolist as consumers do not substitute different varieties at all. In other words, firms do not compete and hence are only affected by their own government's tax.

The same relationship will hold when considering second stage equilibrium taxes, with essentially two groups of players. Signatories' taxes influence non-signatories' quantities negatively and vice versa, except for $\gamma = 0$. Hence, for instance, if governments in signatory countries

want to boost their firms profits by subsidizing their firms, this will automatically reduce foreign firms' quantities. However, if they decide to tax their firms to reduce total output in order to enforce a cartel solution, then this objective is only partially achieved because foreign firm's output will increase. A similar conflict occurs if signatories tax their firms to reduce environmental damages because foreign quantities and hence emissions will increase. Only for $\gamma = 0$ this strategic interaction breaks down.

Second Stage

In this section, we derive equilibrium taxes in the second stage. In order to analyze the importance of each welfare component on equilibrium taxes, the driving forces in our model, the relation to other models in the literature, as well as the impact on the size of stable agreements in the first stage (see the following subsection), we choose a didactically motivated approach and consider three different objective functions. The objective function which corresponds to our welfare function (1) is the last function listed below. The other objective functions comprise only a subset of welfare components. In the political economy literature, these objective functions could be viewed as different political support functions. If they do not comprise all welfare components, this would simply imply that not all interest groups are recognized by governments.

1. $W_i^1 = PS_i + TR_i + CS_i$
2. $W_i^2 = PS_i + TR_i - D_i$
3. $W_i^3 = PS_i + TR_i + CS_i - D_i$

The first objective function ignores environmental damages. The second objective function ignores the consumer surplus and the third objective function is identical to our welfare function (1) as mentioned earlier. For analytic tractability, we henceforth consider two parameter values of γ , namely the "no TFV" scenario with $\gamma = 1$, and the "maximum TFV" scenario with $\gamma = 0$. Despite this simplification, the first-order conditions are huge terms, which cannot be displayed here. Therefore, we provide equilibrium taxes for each objective function in Appendix 1 and provide full details of the first-order conditions upon request. This procedure also applies to the further analysis of equilibrium quantities

once equilibrium taxes have been substituted in stage 3 and equilibrium coalition sizes are determined in stage 1 (see the Appendix).

Inserting equilibrium taxes into equilibrium quantities reveals that we need to impose non-negativity constraints on parameter values in order to ensure positive outputs. This ensures interior solutions, which are easier to analyze than if we allowed for boundary solutions. Essentially, these constraints establish an upper bound on the level of taxes and boil down to requesting that the demand parameter a is larger than marginal production cost c plus a multiple of marginal damages. The exact constraints are stated in Appendix 3, which henceforth are assumed to hold.

We now consider signatories' and non-signatories' taxes across the different objective functions. This will illustrate how different welfare components affect equilibrium taxes. We denote signatories' equilibrium taxes of objective function 1 by $t_i^*(PS, TR, CS)$, objective function 2 by $t_i^*(PS, TR, D)$ and of objective function 3 by $t_i^*(PS, TR, CS, D)$, and the same applies for non-signatories' equilibrium taxes.

Proposition 2 (Comparing Equilibrium Taxes Across Different Objective Functions).

Assume some coalition with m signatories has formed in the first stage and let $n > m > 2$.

Signatories' taxes:

- $t_i^*(PS, TR, D) > t_i^*(PS, TR, CS, D) > t_i^*(PS, TR, CS)$ for $\gamma = \{0, 1\}$.

Non-signatories' taxes:

For $\gamma = 1$:

- $t_j^*(PS, TR, D) < t_j^*(PS, TR, CS, D) < t_j^*(PS, TR, CS)$.

For $\gamma = 0$:

- $t_j^*(PS, TR, D) > t_j^*(PS, TR, CS, D) > t_j^*(PS, TR, CS)$.

Proof. See Appendix 4. □

We first note that signatories' equilibrium taxes are lowered when consumers enter governments' objective function and are increased when

instead damages are considered by governments. The reason is that the consumer surplus is negatively affected by taxes whereas damages are reduced through taxes. Hence, in terms of equilibrium taxes, consumers call for lower and damages for higher equilibrium taxes. In our model, the larger the damage parameter δ compared to the demand parameter a the higher will be the tax and vice versa.

For non-signatories, we observe the same ranking as for signatories if $\gamma = 0$ because then the strategic interaction among firms vanishes. As shown in Proposition 1, if $\gamma = 0$, quantities only depend on own taxes. In contrast for $\gamma = 1$, the strategic interaction among firms is at its maximum and hence also among governments. The ranking of equilibrium taxes for the different objective functions of non-signatories is now reversed to those of signatories. For instance, adding damages to objective function 1 leads to lower equilibrium taxes for non-signatories, already indicating the strategic interaction among signatories and non-signatories, that is, non-signatories free-ride on signatories' emission reduction efforts. This is one version of the free-rider behavior of non-signatories undermining the formation of large stable coalitions which will be analyzed in more detail below.

We now turn to comparing signatories' and non-signatories' taxes for each objective function, which gives rise to further insights into the strategic interaction among signatories and non-signatories.

Proposition 3 (Comparing Equilibrium Taxes for a Given Objective Function).

Objective function 1: $W_i^1 = PS_i + TR_i + CS_i$

- For $\gamma = 1$: $t_i^* = t_j^* < 0$; $\frac{\partial t_i^*}{\partial m} = 0$ and $\frac{\partial t_j^*}{\partial m} = 0$.
- For $\gamma = 0$: $t_i^* < t_j^* < 0$; $\frac{\partial t_i^*}{\partial m} < 0$ and $\frac{\partial t_j^*}{\partial m} = 0$.

Objective Function 2: $W_i^2 = PS_i + TR_i - D_i$

- For $\gamma = 1$: $t_i^* > t_j^*$; $t_i^* >, \leq 0$; $t_j^* < 0$; $\frac{\partial t_i^*}{\partial m} > 0$ and $\frac{\partial t_j^*}{\partial m} < 0$.
- For $\gamma = 0$: $t_i^* > t_j^* > 0$; $\frac{\partial t_i^*}{\partial m} > 0$ and $\frac{\partial t_j^*}{\partial m} = 0$.

Objective Function 3: $W_i^3 = PS_i + TR_i + CS_i - D_i$

- For $\gamma = 1$: $t_i^* > t_j^*$; $t_i^* >, \leq 0$; $t_j^* < 0$; $\frac{\partial t_i^*}{\partial m} > 0$ and $\frac{\partial t_j^*}{\partial m} < 0$.

- For $\gamma = 0$: $t_i^* > t_j^* > 0$ and $\frac{\partial t_i^*}{\partial m} > 0$ if $\delta m + c \leq a < 2n\delta + c$;
 $t_i^* \leq t_j^* \leq 0$ and $\frac{\partial t_i^*}{\partial m} \leq 0$ if $a \geq 2\delta n + c$; $\frac{\partial t_j^*}{\partial m} = 0$.

Taken together⁷:

Objective Functions	Direction of change		Strategic interaction	
	$\gamma = 1$	$\gamma = 0$	$\gamma = 1$	$\gamma = 0$
$PS_i + TR_i + CS_i$	0 , 0	$\downarrow$, 0	<i>independent</i>	<i>independent</i>
$PS_i + TR_i - D_i$	$\uparrow$, $\downarrow$	$\uparrow$, 0	<i>substitutes</i>	<i>independent</i>
$PS_i + TR_i + CS_i - D_i$	$\uparrow$, $\downarrow$	$\uparrow$, 0	<i>substitutes</i>	<i>independent</i>

Proof. See Appendix 1,2,3, and 5. □

There are at least two interesting aspects in Proposition 3. The first aspect relates to the comparison between signatories' and non-signatories' equilibrium taxes where the former internalize externalities within their group. The second aspect relates to the strategic interaction between signatories' and non-signatories' taxes when the coalition is enlarged.

For the first objective function without environmental damages, signatories and non-signatories subsidize production. In principle, there are two externalities for this objective function in this imperfect competition model. Firms produce more than the monopoly quantity (at least if $\gamma \neq 0$), but consumers would like to consume more than the Cournot–Nash quantity and certainly more than the monopoly quantity. Thus, consumers call for low taxes, even subsidies, and from signatory governments' perspective firms' revenues call for high taxes in order to enforce a cartel solution with lower quantities. Though taxes reduce firms' profits, the government collects these taxes and hence taxes are welfare neutral in this model. This last mechanism is not only different from the Brander and Spencer model but also other strategic environment and trade papers for two reasons. In the Brander and Spencer model

⁷The first column "Direction of change", illustrates how signatories' and non-signatories' taxes change with the coalition size. The first entry (arrow) relates to signatories' taxes and the second to non-signatories' taxes. An entry of **0** for signatories means that there is no need for coordination among players. The second column "Strategic interaction" illustrates the strategic interaction between signatories' and non-signatories' taxes. Intuition would suggest that in terms of forming stable coalitions, for a given objective function, this will be easier if taxes are independent than if they are strategic substitutes.

foreign governments are not allowed to react and hence the government subsidizes production in order to shift rents to domestic firms. Moreover, which also applies to other environment and trade models, which allow foreign governments to react, as long as there is no agreement formation among governments, governments cannot enforce a cartel solution. In our model, signatory governments have this possibility and incentive.

For $\gamma = 1$, in our model, these two externalities are going in opposite directions, just cancel out, which explains why $t_i^* = t_j^*$ and why taxes/subsidies do not change with the size of the coalition. For $\gamma = 0$, each firm acts like a monopolist in its market and hence signatory governments do not need to enforce a cartel solution for their firms. In fact, quantities are too low from the perspective of consumers, which explains why signatory governments would choose a higher subsidy than non-signatory governments, the former internalize the externality for their consumers. In other words, there is a positive externality from subsidizing consumption.

For the second objective function, which ignores consumers but considers environmental damages, signatories choose a higher tax than non-signatories in order to internalize the negative externality stemming from emissions. For $\gamma = 0$ when there is no competition among firms, both governments choose a positive tax, though signatory governments choose a higher tax than non-signatory governments. For $\gamma = 1$, with a maximum of strategic interaction among firms, there is the possibility (depending on the parameter values) that signatory governments may choose a negative tax if the number of signatories is small (see Appendix 1). This is because of the trade-off between enforcing a cartel solution and internalizing environmental damages on the one hand, and the fear of losing competitiveness through higher taxes on the other hand. The latter aspect is particularly relevant if an agreement is small and hence the number of free-riders is large. This free-rider behavior is apparent from the fact that non-signatory governments always subsidize their firms for $\gamma = 1$, this subsidy increases with the size of the coalition, that is, $\frac{\partial t_j^*}{\partial m} < 0$ and subsidies are higher than signatory governments' subsidies if they choose subsidies and not a tax. Thus, in a strategic context, even if environmental damages are considered by governments, competition among firms may imply a subsidy and not a tax in equilibrium, but signatory governments choose a lower subsidy than non-signatory governments and in most cases they tax their firms.

The third objective function, including all welfare components, combines the driving forces described for the first and second objective functions. For $\gamma = 1$, this means that signatories have a higher tax than non-signatories, non-signatories always choose a subsidy and non-signatories may choose a tax or a subsidy, which in the latter case only happens if the agreement is small (see Appendix 1). For $\gamma = 0$, when governments do not need to fear competition for their firms, signatory governments need only to balance consumers' and environmental interests. If damages are sufficiently large (represented by the parameter δ) compared to consumers demand (represented by the parameter a) signatory governments will choose a tax and if this is reversed, they may choose a subsidy. In the former case, signatory governments' taxes increase with the coalition size, in the latter case, their subsidies increase with the coalition size.

The second aspect of Proposition 3 is the strategic interaction between signatories' and non-signatories' taxes. In the case of "maximum TFV" with $\gamma = 0$, signatories' and non-signatories' taxes are strategically independent. Non-signatories' taxes do not change with the coalition size m due to the independence of varieties. Signatories' taxes are either increasing or decreasing with the coalition size depending on the relative strength of the externality which they are internalizing. For objective function 1 (2) signatories' taxes decrease (increase) with the coalition size because of the positive externality on consumers (damages). Objective function 3 combines the effects of objective functions 1 and 2 and hence signatories' taxes decrease with the coalition size if the demand parameter a is sufficiently large compared to marginal production costs and global marginal damages and increase if this relation is reversed. In the case of "no TFV" with $\gamma = 1$, signatories' and non-signatories' taxes are strategic substitutes for most objective functions where signatories' taxes are increasing with the coalition size m whereas non-signatories' taxes are decreasing. The exception is the first objective function because externalities cancel out for $\gamma = 1$.

Properties of the Coalition Game

In this section, we analyze some general properties of our coalition game, which will be helpful in order to understand the normative and positive properties of our coalition game. The positive properties also help to

understand the size of stable agreements in the first stage of our game, which we derive in the subsection next.

We define the following properties to analyze the incentive structure to form coalitions and the associated welfare implications.

For all $n \geq m \geq 2$:

- **Superadditivity:** a coalition game is (strictly) superadditive if:

$$mW_{i \in S}(m) \geq (>)[m-1]W_{i \in S}(m-1) + W_{j \notin S}(m-1)$$

- **Positive Externality:** a coalition game exhibits a (strict) positive externality if:

$$W_{j \notin S}(m) \geq (>)W_{j \notin S}(m-1)$$

- **Full Cohesiveness:** a coalition game is (strictly) fully cohesive if:

$$\begin{aligned} mW_{i \in S}(m) + [n-m]W_{j \notin S}(m) \\ \geq (>)[m-1]W_{i \in S}(m-1) + [n-m+1]W_{j \notin S}(m-1). \end{aligned}$$

Superadditivity provides an incentive to join a coalition whereas the positive externality captures the incentive to free-ride. In terms of forming large stable coalitions, the two properties work in opposite directions and typically for large coalitions the positive externality effect is stronger than the superadditivity effect. Full cohesiveness justifies the search for large stable coalitions, even if the grand coalition is not stable. Essentially, global welfare increases when the coalition is enlarged gradually and obtains its maximum in the grand coalition.

Proposition 4 (Properties of the Coalition Game). *In the coalition game, the properties positive externality and full cohesiveness hold strictly for each of the the three objective functions whenever there is an externality across players. For objective function 1 and $\gamma = 1$ when there is no externality across players, these properties hold weakly.*

For all objective functions, superadditivity holds for $\gamma = 0$ and fails for $\gamma = 1$ whenever there is an externality across players. For $\gamma = 1$ it only holds for the move from a coalition with $n-1$ signatories to the grand coalition with n signatories. For objective function 1 and $\gamma = 1$ when there is no externality across players, superadditivity holds weakly. More specifically⁸:

⁸Note: “+”, holds strictly, “0”, holds weakly; and “−” does not hold generally.

Objective functions	Positive externality		Superadditivity		Full cohesiveness	
	$\gamma = 1$	$\gamma = 0$	$\gamma = 1$	$\gamma = 0$	$\gamma = 1$	$\gamma = 0$
$PS_i + TR_i + CS_i$	0	+	0	+	0	+
$PS_i + TR_i - D_i$	+	+	−	+	+	+
$PS_i + TR_i + CS_i - D_i$	+	+	−	+	+	+

Proof. See Appendix 6. □

Clearly, for objective function 1 and $\gamma = 1$ these properties are meaningless as externalities cancel out. Proposition 4 confirms that the normative property of the game full cohesiveness holds. The larger are, the coalitions, the larger will be global welfare. It also confirms that non-signatories benefit from the enlargement of the coalition via positive externalities. This provides an incentive to free-ride. Interestingly, the incentive to join a coalition, captured by the property superadditivity, is only positive if $\gamma = 0$ but is negative if $\gamma = 1$ and whenever coalition formation would matter (i.e., full cohesiveness strictly holds). In the latter case, signatories’ taxes increase with the coalition size and the reverse is true for non-signatories as shown in Proposition 3. That is, strategies are substitutes, and hence the efforts of signatories are undermined by non-signatories’ reaction. This countervailing or leakage effect renders the enlargement of the coalition unsuccessful. As recently shown in Bayramoglu *et al.* (2016), if the move from a coalition with $m - 1$ to m members is not superadditive, then the coalition with m signatories cannot be internally stable. In other words, superadditivity is a necessary (though not sufficient) condition for internal stability in a positive externality game. Hence, if superadditivity fails for all $m \leq n - 1$ for $\gamma = 1$, we only need to test for stability of the grand coalition. Our overall results are summarized in Proposition 5, which looks at the stability of coalitions in the first stage.

First Stage

In this section, we present the results for the first stage, that is, the stability of coalitions.

Proposition 5 (Coalition Stability). *Let m^* denote the size of an internally and externally stable coalition and let $n > 5$. For the three*

objective functions, whenever there is an incentive for countries to coordinate their policy (i.e., $\frac{\partial t_i^*}{\partial m} \neq 0$), the following results are obtained: For $\gamma = 1$: $m^* = 1$ and for $\gamma = 0$: $m^* = 3$.

More specifically:

Objective functions	IES	
	$\gamma = 1$	$\gamma = 0$
$PS_i + TR_i + CS_i$	-	$m^* = 3$
$PS_i + TR_i - D_i$	$m^* = 1$	$m^* = 3$
$PS_i + TR_i + CS_i - D_i$	$m^* = 1$	$m^* = 3$

Proof. See Appendix 7. □

Proposition 5 shows that if $\gamma = 1$ there are no stable coalitions, except for objective function 1 for which coalition formation is meaningless as externalities across players cancels out. This can be related to two previous results. In Proposition 3, we showed that taxes are strategic substitutes, and in Proposition 4, we showed that this implied that superadditivity generally failed.

In contrast, if $\gamma = 0$ there is at least a stable coalition of three countries. This can also be inferred from our previous results. In Proposition 3, we showed that taxes are strategically independent in these scenarios, and in Proposition 4 we showed that superadditivity holds. However, interestingly, despite superadditivity holds, stable coalitions are small because of the positive externality property.⁹

It is interesting that these results hold for both values of γ irrespective of the weight consumers and damages receive in governments' objective function, which stresses that they are quite robust. Even considering consumers' interests apart from environmental damages in governments' objective function do not mitigate the free-rider incentives sufficiently enough, and hence the establishment of a stable agreement is difficult. Only if the competition among firms is sufficiently reduced through a high taste for variety by consumers, can small agreements be stable (see Footnote 9). Thus, in relating our results to the strategic environment and trade literature without coalition formation, their

⁹It is probably not surprising that for intermediate values of γ between 0 and 1, one finds that the equilibrium coalition size lies between $m^* = 1$ and $m^* = 3$. For instance, for $\gamma = 0.5$ we find $m^* = 2$.

main conclusions remain valid for $\gamma = 1$ as no stable agreement forms when allowing for this possibility. However, if the taste for variety is sufficiently strong, more positive results can be obtained, even though the normative benchmark of the social optimum will not be obtained.

Concluding Remarks

In this paper, we analyzed a strategic trade model in the spirit of Brander and Spencer (1985). We introduced three additional features, which have been considered in the literature, though in isolation. Firstly, consumers matter for governments because goods are not sold to a third market. Moreover, environmental damages matter because production releases a global pollutant. Second, we consider horizontal product differentiation with consumers having a taste for variety (TFV). For analytical tractability, we focused on two extreme assumptions of TFV: no TFV and maximum TFV where the former assumption corresponds to the standard assumption in the literature that goods are perfect substitutes. These two extreme assumptions were sufficient because the driving forces identified would also be present for intermediate cases. Thirdly, we considered the possibility that governments can coordinate their policy by forming coalitions, which, to the best of our knowledge, has not been considered in the literature so far. Policy coordination is related to an emission tax, which is de facto an output tax because of a constant output-emission ratio. In our simple model this tax is efficient as firms have no option to switch to a more environmentally friendly technology. Stability of an agreement was tested by invoking the concept of internally and externally stable cartels.

We demonstrated that the formation of agreements is globally beneficial. Global welfare increases with the size of agreements and obtains its maximum if the grand coalition forms (full cohesiveness). However, the grand coalition or even smaller coalitions may not be stable because of two reasons. Firstly, the benefits from policy coordination are non-exclusive, a features which we related to the property of positive externality of coalition formation. Secondly, the gains from cooperation for those involved in enlarging coalitions may be small or even negative if policy instruments are strategic substitutes. That is, superadditivity may fail.

We showed that for the “no TFV” scenario, signatories of an agreement increase their taxes with the size of the agreement. Signatory governments have an incentive to internalize two negative externalities, both associated with high quantities. A reduction of output stabilizes the price in the output cartel and also reduces environmental damages. Non-signatories free-ride on signatories’ efforts and lower their taxes. Hence, taxes are strategic substitutes between signatories and non-signatories. In our model, this meant that no agreement was stable. In contrast, for the “maximum TFV” scenario, foreign taxes have no effect on domestic firms’ output. In the context of an agreement, this implies that taxes of signatories (non-signatories) have no effect on the output of non-signatories’ (signatories’) firms. We found that this implies that taxes between signatories and non-signatories become strategically independent. Regardless whether signatories increase or decrease their tax with the size of the agreement, non-signatories’ equilibrium taxes do not change. This reduces the free-rider incentive, but it remains positive, which explains that this led only to small stable coalitions.

To our knowledge, this is the first attempt to introduce consumers’ taste for variety to the literature of international environmental agreements and trade. Our stylized model allows for exploring future research avenues in terms of additional policy instruments, like tax border adjustments (Helm and Schmidt, 2015), relaxing the symmetry assumption and further investigations of sub-features of TFV, such as ideal varieties or asymmetric consumers’ TFV between countries.

Appendices

A detailed appendix with the full details of all derivations is available upon request. Below, we summarize the most important steps in the derivation in a compact form for $\gamma = 1$ and $\gamma = 0$.

Appendix 1: Equilibrium Taxes for all Objective Functions

For each function, we derive the FOCs for signatories and non-signatories in stage 2 of the game. Solving these conditions simultaneously, we find the equilibrium taxes for signatories and non-signatories (for which

the signs follow from using the non-negativity constraints derived in Appendix 3).

- For $\gamma = 1$:

$$t_i^*(PS, TR, CS) = t_j^*(PS, TR, CS) = -\frac{a-c}{n} < 0$$

$$t_i^*(PS, TR, D) = \frac{\delta[n(m(n-m+3)-n-1)+m(2-m)]}{m(n(n-m+1)-m+2)} - \frac{(a-c)(n-2m+1)}{m(n(n-m+1)-m+2)} > 0, \leq 0$$

$$t_j^*(PS, TR, D) = -\frac{\delta(m-2)(n+1)+(a-c)(n-1)}{n(n-m+1)-m+2} < 0$$

$$t_i^*(PS, TR, CS, D) = \frac{n\delta[m(2-m)+n(n(m-1)+m(3-m)-1)]}{mn(n^2+(1-m)(n+1))} - \frac{m(a-c)(n^2+(1-m)(n+1))}{mn(n^2+(1-m)(n+1))} > 0, \leq 0$$

$$t_j^*(PS, TR, CS, D) = -\frac{n\delta(m-2)(n+1)+(a-c)}{n(n^2+(1-m)(n+1))} \frac{(n^2+(1-m)(n+1))}{n(n^2+(1-m)(n+1))} < 0$$

- For $\gamma = 0$:

$$t_i^*(PS, TR, CS) = -\frac{(a-c)m}{2n-m} < 0$$

$$t_j^*(PS, TR, CS) = -\frac{a-c}{2n-1} < 0$$

$$t_i^*(PS, TR, D) = \delta m > 0$$

$$t_j^*(PS, TR, D) = \delta > 0$$

$$t_i^*(PS, TR, CS, D) = \frac{(2n\delta - a + c)m}{2n-m} > 0, \leq 0$$

$$t_j^*(PS, TR, CS, D) = \frac{2n\delta - a + c}{2n-1} > 0, \leq 0$$

Appendix 2: Definitions

There are certain terms that repeatedly show up in the following. They are listed below.

$$\Psi_1 = n(n - m + 1) - m + 2 = n^2 - nm + n - m + 2$$

$$\Psi_2 = n^2 + (1 - m)(n + 1) = n^2 - nm + n - m + 1$$

$$\begin{aligned}\Psi_3 &= n^2(m - 1) - n(m - 1)^2 - m(m - 2) \\ &= n^2m - n^2 - nm^2 + 2nm - n - m^2 + 2m\end{aligned}$$

$$\Psi_4 = n(m - 1) - m(m - 2) = nm - n - m^2 + 2m$$

$$\Psi_5 = (m - 2)(n + 1) = mn + m - 2n - 2$$

$$\Psi_6 = (m - 1)(n + 1) = mn + m - n - 1$$

$$\begin{aligned}\Psi_7 &= (n - m)(n - m + 1) - m(1 - m) + 2 \\ &= n^2 - 2nm + 2m^2 + n - 2m + 2\end{aligned}$$

$$\begin{aligned}\Psi_8 &= n(n^2 + n + 1) - 2nm(n - m + 1) + m^2 \\ &= n^3 - 2n^2m + 2nm^2 + n^2 - 2nm + m^2 + n\end{aligned}$$

It can be shown that all $\Psi_k > 0$, $\forall n$ and $\forall m \leq n$.

Appendix 3: Non-negativity Constraints

Inserting equilibrium taxes into equilibrium output levels, gives the quantities below, from which it is evident that for the first objective function no non-negativity constraint needs to be imposed apart from $a > c$. For the second and third objective functions, additional conditions need to be imposed as explained below.

- For $\gamma = 1$:

$$\begin{aligned}q_{.i}^*(\text{PS}, \text{TR}, \text{CS}) &= q_{.j}^*(\text{CS}, \text{PS}, \text{TR}) = \frac{a - c}{n} \\ q_{.i}^*(\text{PS}, \text{TR}, D) &= \frac{(a - c)(n - m + 1) - \delta\Psi_3}{m\Psi_1} \\ q_{.j}^*(\text{PS}, \text{TR}, D) &= \frac{n(a - c) + \delta(n(m - 1) + m - 2)}{\Psi_1}\end{aligned}$$

$$q_{.i}^*(\text{PS}, \text{TR}, \text{CS}, D) = \frac{(a-c)}{n} - \frac{\Psi_3 \delta}{m\Psi_2}$$

$$q_{.j}^*(\text{PS}, \text{TR}, \text{CS}, D) = \frac{(a-c)}{n} + \frac{\delta n(n(m-1) + m - 2)}{n\Psi_2}$$

- For $\gamma = 0$:

$$q_{.i}^*(\text{PS}, \text{TR}, \text{CS}) = \frac{n(a-c)}{2n-m}$$

$$q_{.j}^*(\text{PS}, \text{TR}, \text{CS}) = \frac{n(a-c)}{2n-1}$$

$$q_{.i}^*(\text{PS}, \text{TR}, D) = \frac{a-c-\delta m}{2}$$

$$q_{.j}^*(\text{PS}, \text{TR}, D) = \frac{a-c-\delta}{2}$$

$$q_{.i}^*(\text{PS}, \text{TR}, \text{CS}, D) = \frac{n(a-c-\delta m)}{2n-m}$$

$$q_{.j}^*(\text{PS}, \text{TR}, \text{CS}, D) = \frac{n(a-c-\delta)}{2n-1}$$

For the second and third objective functions, the following non-negativity constraints need to be imposed.

- Objective Function 2: $W_i^2 = \text{PS}_i + \text{TR}_i - D_i$
 - For $\gamma = 1$ signatories' non-negativity constraint is given by $a > \tilde{a}_1 = \frac{\delta\Psi_3}{n-m+1} + c$, and for non-signatories by $a > c$, with $\tilde{a}_1 > c$.
 - For $\gamma = 0$ signatories' non-negativity constraint is given by $a > \tilde{a}_2 = \delta m + c$, and for non-signatories by $a > \tilde{a}_3 = \delta + c$, with $\tilde{a}_2 > \tilde{a}_3$.
- Objective Function 3: $W_i^3 = \text{PS}_i + \text{TR}_i + \text{CS}_i - D_i$
 - For $\gamma = 1$ signatories' non-negativity constraint is given by $a > \tilde{a}_4 = \frac{\delta n\Psi_3}{m\Psi_2} + c$, and for non-signatories by $a > c$, with $\tilde{a}_4 > \tilde{a}_2$.

- For $\gamma = 0$ the non-negativity constraints are the same as for the second objective function above ($a > \tilde{a}_2$ for signatories and $a > \tilde{a}_3$ for non-signatories).

It is straightforward to show that $\tilde{a}_1 > \tilde{a}_4$. Throughout the whole paper, we assume the most restrictive constraint to hold for comparison for a particular objective function and across objective functions, noting that $n \geq m \geq 2$.

Appendix 4: Comparing Equilibrium Taxes Across Different Objective Functions

Assume $n \geq m > 2$ and the appropriate non-negativity constraints in Appendix 3 to hold. Then, using equilibrium taxes in Appendix 1, and the definitions in Appendix 2, we find:

For $\gamma = 1$:

$$\begin{aligned} t_i^*(\text{PS}, \text{TR}, D) - t_i^*(\text{PS}, \text{TR}, \text{CS}, D) &= \frac{(n+1)\Psi_4}{m\Psi_1} \left(\frac{a-c}{n} - \frac{\delta}{\Psi_2} \right) > 0 \\ t_j^*(\text{PS}, \text{TR}, D) - t_j^*(\text{PS}, \text{TR}, \text{CS}, D) &= -\frac{(m-2)(n+1)}{\Psi_1} \cdot \left(\frac{a-c}{n} - \frac{\delta}{\Psi_2} \right) < 0 \\ t_i^*(\text{PS}, \text{TR}, \text{CS}) - t_i^*(\text{PS}, \text{TR}, \text{CS}, D) &= -\frac{\delta(n+1)\Psi_4}{m\Psi_2} < 0 \\ t_j^*(\text{PS}, \text{TR}, \text{CS}) - t_j^*(\text{PS}, \text{TR}, \text{CS}, D) &= \frac{\delta\Psi_5}{\Psi_2} > 0 \end{aligned}$$

For $\gamma = 0$:

$$\begin{aligned} t_i^*(\text{PS}, \text{TR}, D) - t_i^*(\text{PS}, \text{TR}, \text{CS}, D) &= \frac{m(a-c-\delta m)}{2n-m} > 0 \\ t_j^*(\text{PS}, \text{TR}, D) - t_j^*(\text{PS}, \text{TR}, \text{CS}, D) &= \frac{a-c-\delta}{2n-1} > 0 \\ t_i^*(\text{PS}, \text{TR}, \text{CS}) - t_i^*(\text{PS}, \text{TR}, \text{CS}, D) &= -\frac{2nm\delta}{2n-m} < 0 \\ t_j^*(\text{PS}, \text{TR}, \text{CS}) - t_j^*(\text{PS}, \text{TR}, \text{CS}, D) &= -\frac{2n\delta}{2n-1} < 0 \end{aligned}$$

Appendix 5: Comparing Equilibrium Taxes for Each Objective Function

Using equilibrium taxes as listed in Appendix 1, and the definitions in Appendix 2, we find:

For $\gamma = 1$:

$$t_i^*(\text{PS, TR, CS}) - t_j^*(\text{PS, TR, CS}) = 0$$

$$t_i^*(\text{PS, TR, D}) - t_j^*(\text{PS, TR, D}) = \frac{\Psi_6(n\delta + a - c)}{m\Psi_1} > 0$$

$$t_i^*(\text{PS, TR, CS, D}) - t_j^*(\text{PS, TR, CS, D}) = \frac{n\delta\Psi_6}{m\Psi_2} > 0$$

For $\gamma = 0$:

$$t_i^*(\text{PS, TR, CS}) - t_j^*(\text{PS, TR, CS}) = -\frac{2n(a - c)(m - 1)}{(2n - m)(2n - 1)} < 0$$

$$t_i^*(\text{PS, TR, D}) - t_j^*(\text{PS, TR, D}) = \delta(m - 1) > 0$$

$$t_i^*(\text{PS, TR, CS, D}) - t_j^*(\text{PS, TR, CS, D}) = \frac{2n(2n\delta - a + c)(m - 1)}{(2n - m)(2n - 1)}$$

which is positive if $\delta m + c < a \leq 2n\delta + c$ (where $\delta m + c < a$ is the non-negativity constraint) and negative if $a > 2n\delta + c$.

Furthermore:

For $\gamma = 1$:

$$\frac{\partial t_i^*(\text{PS, TR, CS})}{\partial m} = 0$$

$$\frac{\partial t_j^*(\text{PS, TR, CS})}{\partial m} = 0$$

$$\frac{\partial t_i^*(\text{PS, TR, D})}{\partial m} = \frac{(n + 1)(n\delta + a - c)\Psi_7}{m^2\Psi_1^2} > 0$$

$$\frac{\partial t_j^*(\text{PS, TR, D})}{\partial m} = -\frac{(n + 1)(n - 1)(n\delta + a - c)}{\Psi_1^2} < 0$$

$$\frac{\partial t_i^*(\text{PS}, \text{TR}, \text{CS}, D)}{\partial m} = \frac{\delta(n+1)\Psi_8}{m^2\Psi_2^2} > 0$$

$$\frac{\partial t_j^*(\text{PS}, \text{TR}, \text{CS}, D)}{\partial m} = -\frac{\delta(n+1)(n^2 - n - 1)}{\Psi_2^2} < 0$$

For $\gamma = 0$:

$$\frac{\partial t_i^*(\text{PS}, \text{TR}, \text{CS})}{\partial m} = -\frac{2n(a-c)}{(2n-m)^2} < 0$$

$$\frac{\partial t_j^*(\text{PS}, \text{TR}, \text{CS})}{\partial m} = 0$$

$$\frac{\partial t_i^*(\text{PS}, \text{TR}, D)}{\partial m} = \delta > 0$$

$$\frac{\partial t_j^*(\text{PS}, \text{TR}, D)}{\partial m} = 0$$

$$\frac{\partial t_i^*(\text{PS}, \text{TR}, \text{CS}, D)}{\partial m} = \frac{2n(2n\delta - a + c)}{(2n-m)^2}$$

which is positive if $\delta m + c < a \leq 2n\delta + c$ (where $\delta m + c < a$ is the non-negativity constraint) and negative if $a > 2n\delta + c$.

$$\frac{\partial t_j^*(\text{PS}, \text{TR}, \text{CS}, D)}{\partial m} = 0$$

Appendix 6: Properties of the Coalition Game

Objective Function 1¹⁰: $W_i^1 = \text{PS}_i + \text{TR}_i + \text{CS}_i$:

- For $\gamma = 1$:

$$\text{PEP} = 0$$

$$\text{SAD} = 0$$

$$\text{FC} = 0$$

¹⁰PEP, positive externality property; SAD, superadditivity; FC, full cohesiveness.

- For $\gamma = 0$:

$$\begin{aligned} \text{PEP} &= \frac{(32n^3 + m(n(-28n + 8m) - m^2 + m - 1))}{(2n - 1)^2(2n - m + 1)^2(2n - m)^2} \\ &\quad \cdot \frac{(m - 1)n^2(a - c)^2}{2} > 0 \\ \text{SAD} &= \frac{(4n^2m - 4nm^2 + m^3 + 8nm - 3m^2 - 8n + 3m)}{(2n - 1)^2(2n - m + 1)^2(2n - m)} \\ &\quad \cdot \frac{(m - 1)n^2(a - c)^2}{2} > 0 \end{aligned}$$

FC follows from PEP and SAD.

Objective Function 2: $W_i^2 = \text{PS}_i + \text{TR}_i - D_i$:

- For $\gamma = 1$:

$$\text{PEP} = \frac{(n\delta + a - c)^2(2n^2 - 2nm + 3n - 2m + 5)(n + 1)n^2}{\Psi_1^2(n^2 - nm + 2n - m + 3)^2} > 0$$

$$\begin{aligned} \text{SAD} &= -\frac{(n\delta + a - c)^2(n^4 - 2n^3m + n^2m^2 + nm)(n + 1)n}{\Psi_1^2(n^2 - nm + 2n - m + 3)^2} \\ &\quad - \frac{(n\delta + a - c)^2(-m^2 - 4n + 3m - 1)(n + 1)n}{\Psi_1^2(n^2 - nm + 2n - m + 3)^2} > 0, \end{aligned}$$

$$m = n, \quad \text{and} \quad < 0 \quad \forall m < n$$

$$\text{FC} = \frac{(n\delta + a - c)^2(n + 1)^2n\Omega}{\Psi_1^2(n^2 - nm + 2n - m + 3)^2} > 0$$

$$\Omega := n^3 - 2n^2m + nm^2 + 2n^2 - 3nm + m^2 + 3n - 3m + 1$$

- For $\gamma = 0$:

$$\text{PEP} = n(m - 1)\delta^2 > 0$$

$$\text{SAD} = \frac{1}{4}n(m - 1)m\delta^2 > 0$$

FC follows from PEP and SAD.

Objective Function 3: $W_i^3 = \text{PS}_i + \text{TR}_i + \text{CS}_i - D_i$:

- For $\gamma = 1$:

$$\text{PEP} = \frac{1}{2} \frac{\delta^2(2n^2 - 1)(2n^2 - 2nm + 3n - 2m + 3)(n + 1)n^2}{(n^2 - \Psi_5)^2 \Psi_2^2} > 0$$

$$\begin{aligned} \text{SAD} = & -\frac{1}{2} \frac{\delta^2(2n^5 - 4n^4m + 2n^3m^2 - 4n^3 + 8n^2m)(n + 1)n^2}{(n^2 - \Psi_5)^2 \Psi_2^2} \\ & - \frac{1}{2} \frac{\delta^2(-4nm^2 - 8n^2 + 9nm - 2m^2 - 4n + 3m)(n + 1)n^2}{(n^2 - \Psi_5)^2 \Psi_2^2} > 0, \end{aligned}$$

$$m = n, \quad < 0, \quad \text{and } \forall m < n$$

$$\text{FC} = \frac{1}{2} \frac{\delta^2(n + 1)^2 n^3 \Phi}{(n^2 - \Psi_5)^2 \Psi_2^2} > 0$$

$$\Phi := 2n^3 - 4n^2m + 2nm^2 + 4n^2 - 6nm + 2m^2 + 4n - 4m + 1$$

- For $\gamma = 0$:

$$\text{PEP} = \frac{1}{2} \frac{(n\delta - a + c)^2(m - 1)n^2 \Upsilon}{(2n - 1)^2(2n - m + 1)^2(2n - m)^2} > 0$$

$$\Upsilon := 32n^3 - 28n^2m + 8nm^2 - m^3 + m^2 - m$$

$$\text{SAD} = \frac{1}{2} \frac{(n\delta - a + c)^2(m - 1)n^2 \Xi}{(2n - 1)^2(2n - m + 1)^2(2n - m)^2} > 0$$

$$\Xi := 4n^2m - 4nm^2 + m^3 + 8nm - 3m^2 - 8n + 3m$$

FC follows from PEP and SAD.

Coalition Stability

Objective Function 1: $W_i^1 = \text{PS}_i + \text{TR}_i + \text{CS}_i$:

- For $\gamma = 1$:

$$W_i(S) - W_i(S \setminus \{i\}) = 0$$

- For $\gamma = 0$:

$$\begin{aligned} W_i(S) - W_j(S \setminus \{i\}) \\ = -\frac{1}{2} \frac{n^2(a-c)^2(m-1)(2nm-m^2-6n+3m-1)}{(2n-1)(2n-m+1)^2(2n-m)} > 0, \\ \forall m \leq 3, \text{ and } < 0, \forall m > 3 \end{aligned}$$

Objective Function 2: $W_i^2 = \text{PS}_i + \text{TR}_i - D_i$:

- For $\gamma = 1$

$$W_i(S) - W_j(S \setminus \{i\}) = -\frac{n(n\delta + a - c)^2(n+1)(\Lambda - \Theta - \Gamma)}{m\Psi_1^2(n^2 - nm + 2n - m + 3)^2} < 0$$

$$\begin{aligned} \Lambda &:= n^4m - 2n^3m^2 + n^2m^3 - n^4 + 4n^3m \\ \Theta &:= -5n^2m^2 + 2nm^3 - 4n^3 + 13n^2m - 10nm^2 \\ \Gamma &:= m^3 - 10n^2 + 17nm - 7m^2 - 12n + 15m - 9 \end{aligned}$$

- For $\gamma = 0$:

$$\begin{aligned} W_i(S) - W_j(S \setminus \{i\}) &= -\frac{1}{4}n\delta^2(m-1)(m-3) \geq 0, \\ \forall m \leq 3, \text{ and } < 0, \forall m > 3 \end{aligned}$$

Objective Function 3: $W_i^3 = \text{PS}_i + \text{TR}_i + \text{CS}_i - D_i$:

- For $\gamma = 1$:

$$W_i(S) - W_j(S \setminus \{i\}) = -\frac{1}{2} \frac{\delta^2 n^2 (n+1)(\Lambda - \Theta - \Gamma)}{m\Psi_2^2(n^2 - \Psi_5)^2} < 0$$

$$\begin{aligned} \Lambda &:= 2n^5m - 4n^4m^2 + 2n^3m^3 - 2n^5 + 8n^4m - 10n^3m^2 \\ \Theta &:= 4n^2m^3 - 8n^4 + 22n^3m - 16n^2m^2 + 2nm^3 - 16n^3 \\ \Gamma &:= 28n^2m - 12nm^2 - 16n^2 + 19nm - 2m^2 - 8n + 3m \end{aligned}$$

- For $\gamma = 0$:

$$\begin{aligned} W_i(S) - W_j(S \setminus \{i\}) \\ = -\frac{1}{2} \frac{n^2(2n\delta - a + c)^2(m-1)(2nm-m^2-6n+3m-1)}{(2n-1)(2n-m+1)^2(2n-m)} > 0, \\ \forall m \leq 3, \text{ and } < 0, \forall m > 3. \end{aligned}$$

References

- Appleyard, D. and A. Field (2014), *International Economics*, 8th Edition, International, New York, NY 10020: McGraw-Hill.
- Barrett, S. (1994a), "Self-Enforcing International Environmental Agreements", *Oxford Economic Papers*, 46, 878–94.
- Barrett, S. (1994b), "Strategic Environmental Policy and International Trade", *Journal of Public Economics*, 54, 325–38.
- Barrett, S. (2002), "Consensus Treaties", *Journal of Institutional and Theoretical Economics*, 158(4), 529–47.
- Barrett, S. (2006), "Climate Treaties and "Breakthrough" Technologies", *American Economic Review*, 96(2), 22–5.
- Bayramoglu, B., M. Finus, and J.-F. Jacques (2016), "Climate Coalitions in Mitigation-Adaptation Game", Bath Economic Research Papers, No. 51/16.
- Brander, J. A. and B. J. Spencer (1985), "Export Subsidies and International Market Share Rivalry", *Journal of International Economics*, 18, 83–100.
- Conrad, K. (1993), "Taxes and Subsidies for Pollution-Intensive Industries as Trade Policy", *Journal of Environmental Economics and Management*, 25, 121–35.
- Copeland, B. R. and M. S. Taylor (2005), "Free Trade and Global Warming: A Trade Theory View of the Kyoto Protocol", *Journal of Environmental Economics and Management*, 49, 205–34.
- d'Aspremont, C., A. Jacquemin, J. J. Gabszewicz, and J. A. Weymark (1983), "On the Stability of Collusive Price Leadership", *Canadian Journal of Economics*, 16(1), 17–25.
- Di Comite, F., J.-F. Thisse, and H. Vandenbussche (2014), "Vertical Differentiation in Export Markets", *Journal of International Economics*, 93(1), 50–66.
- Dixit, A. K. and J. Stiglitz (1977), "Monopolistic Competition and Optimum Product Diversity", *The American Economic Review*, 67(3), 297–308.
- Eichner, T. and R. Pethig (2012), "Stable Climate Coalitions (Nash) and International Trade. Center for Economic Studies and Ifo Institute", CESifo Working Papers, No. 3915.

- Eichner, T. and R. Pethig (2013), "Self-Enforcing Environmental Agreements and International Trade", *Journal of Public Economics*, 102, 37–50.
- Eichner, T. and R. Pethig (2014), "Global Environmental Agreements and International Trade: Asymmetry of Countries Matters. Center for Economic Studies and Ifo Institute", CESifo Working Papers, No. 5051.
- Eichner, T. and R. Pethig (2015), "Self-Enforcing International Environmental Agreements and Trade: Taxes Versus Caps", *Oxford Economic Papers*, 67, 897–917.
- El-Sayed, A. and S. Rubio (2014), "Sharing R&D Investments in Cleaner Technologies to Mitigate Climate Change", *Resource and Energy Economics*, 38, 168–80.
- Finus, M. and A. Caparrós (2015), "Introduction", in *Handbook on Game Theory and International Environmental Cooperation: Essential Readings*, ed. M. Finus and A. Caparrós, Cheltenham, UK: Edward Elgar, xvii–xliv.
- Finus, M. and S. Maus (2008), "Modesty May Pay!", *Journal of Public Economic Theory*, 10(5), 801–26.
- Helm, C. and R. C. Schmidt (2015), "Climate Cooperation with Technology Investments and Border Carbon Adjustment", *European Economic Review*, 75, 112–30.
- Helpman, E. and P. R. Krugman (1985), *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*, Cambridge, Massachusetts. London, UK: The MIT Press.
- Hoel, M. and A. de Zeeuw (2010), "Can a Focus on Breakthrough Technologies Improve the Performance of International Environmental Agreements?", *Environmental and Resource Economics*, 47, 395–406.
- Kennedy, P. (1994), "Equilibrium Pollution Taxes in Open Economies with Imperfect Competition", *Journal of Environmental Economics and Management*, 27, 49–63.
- Loke, W.-H. and L. A. Winters (2012), "Why is Open Regionalism Not Always Good?", *Journal of Economic Integration*, 27(3), 437–65.
- Ulph, A. (1996a), "Environmental Policy and International Trade when Governments and Producers Act Strategically", *Journal of Environmental Economics and Management*, 30, 265–81.

- Ulph, A. (1996b), “Environmental Policy Instrument and Imperfectly Competitive International Trade”, *Environmental and Resource Economics*, 7, 333–55.
- UNFCCC (2015), “Synthesis Report on the Aggregate Effect of the Intended Nationally Determined Contributions”, United Nations Framework Convention on Climate Change. FCCC/CP/2015/7.
- Yi, S.-S. (1996), “Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism Is Good”, *Journal of International Economics*, 41, 153–77.
- Yi, S.-S. (2000), “Free-Trade Areas and Welfare: An Equilibrium Analysis”, *Review of International Economics*, 8(2), 336–47.