Pushing the Tipping in International Environmental Agreements

Lorenzo Cerda Planas*

Last revision: May 31^{st}, 2015

Abstract

This paper intends to provide an alternative approach to the formation of International Environmental Agreements (IEA). The existing consensus within the literature is that there are either too few signatories or that the emissions of signatories are almost the same as business as usual (BAU). I start from a well-known model (Barrett 1997), adding heterogeneity in countries' marginal abatement costs (low and high) and in damages suffered (or corresponding environmental concern). I also allow for technological transfers and border taxes. I show that, depending on the parameter conditions, using either mechanism one at a time or both together induces the grand and abating coalition. I analyse if a retaliation tax is profitable for non-signatories and in which conditions this is the case. Finally, if transfers are part of the recipe, I show that it is optimal to make these transfers to the least-green countries.

JEL Classification: F53, C63, C72, F18, O32.
Keywords: Self-enforcing environmental agreements, technological transfers, border tax, tipping.

*Paris School of Economics, C.E.S. - Université de Paris 1 Panthéon-Sorbonne; 106-112, Boulevard de l'Hôpital, 75647 Paris Cedex 13, France. Email: lorenzo.cerda@gmail.com

Acknowledgments: I want to thank Bertrand Wigniolle and Scott Barrett for their invaluable help and support. I would also like to thank Lisa Anouliès for her thorough revision.

1 Introduction

The present paper focuses on the idea that the creation of an International Environmental Agreement (IEA) that brings together many countries can be a colossal task to achieve. Barrett (1994), Barrett (2003), Rubio and Ulph (2006), and Eichner and Pethig (2013), among others, show that the number of signatories of self-enforcing IEAs generally does not exceed three or four; or if it does, their emissions are almost the same as business as usual (BAU). Recent failures to reach an effective agreement concerning climate change are a clear example of this fact. Some solutions that have been analysed to arrive at a successful IEA involve the idea of starting with a coalition composed of a small sub-set of countries and then incorporating more countries. A first possibility explored by Heal and Kunreuther (2011) suggests enlarging a small coalition in a cascading process. They assume a positive reinforcing effect among signatories, which induces the accession of more countries. In a similar manner, Barrett (2006) shows that if there is a green technology that exhibits increasing returns to adoption, a tipping coalition size exists after which countries join to finally reach the grand coalition. A second line of thought is analysed by Carraro and Siniscalco (1993), Hoel and Schneider (1997), and Barrett (2001), where transfers are used in order to induce more countries to join the IEA. The first paper shows that a commitment problem prevents the formation of the grand coalition – or if it forms, the countries still function as though it were business as usual (BAU). The commitment problem comes from the fact that the IEA set-up resembles a chicken game: Signatories are better off compared to not having an IEA at all, but they would prefer to have others to sign and abstain themselves. Hence, when transfers are made in order to enlarge the coalition, the original signatories prefer to leave. Hoel & Schneider (1997) show a similar effect, where commitment is gained by conformity. Barrett (2001) finds a similar result as Carraro & Siniscalco (1993) but with asymmetric countries. If asymmetry is weak, the commitment issue arises. If there is strong asymmetry, transfers can sustain a superior outcome compared to an IEA without side payments.

Given all these barriers, I explore a new tactic. I show that using a border tax, a technological transfer, or both can effectively induce and sustain the grand and meaningful coalition.(1) I find the conditions under which each case is optimal. It turns out that there is no general recipe, although at least one instrument is always better than nothing. I also show that if transfers are part of the optimal solution, they have to be made to the least-green countries. This choice improves the chances of success while minimizing the amount of transfers. These results could be quite important in face of the recent failures in reaching an IEA that effectively tackles climate change. It is possible that the reasonings

(1)Meaningful in the sense that countries are actually making an effort in abating and not simply acting in a way similar to BAU.

presented in this paper could light the ongoing negotiations and offer potential solutions.

In order to develop this idea, I build a model that is an expanded version of Barrett (1997). In his paper, the world is composed of identical countries, each containing one firm that produces an homogeneous and traded good and pollution as a side product. In Barrett's three-stage game, each country decides to join an IEA or not (one shot). In the second stage, the formed coalition is group rational and plays accordingly, with non-signatories playing as singletons. Finally, firms maximize their profits by choosing their segmented outputs à la Cournot, and markets clear.

Starting from this set-up, I first add asymmetries among countries in terms of damages from emissions and abatement costs. These asymmetries are empirically motivated. Societies exhibit green behaviour, even though direct 'rational' thinking could command otherwise. This could be related to moral reasons rather than economic ones. As shown in Cerda (2015), structurally similar countries (in terms of development level and political system) can end up behaving differently with respect to global environment. This could be for historical reasons, but the point is that some countries have become (quite) aware of the environmental problem at hand and have started acting accordingly. Hence, differences among countries are how green they are and what abatement technology they have in place. With respect to the first point, an equivalent interpretation could be how countries are affected by pollution. Therefore, having higher marginal damages coming from emission will be treated as a synonym of being greener. On the abatement side, countries can either have bad technology (high marginal cost of abatement) or good technology (low marginal cost of abatement). Conversely, it can be thought that historically green countries have invested in cleaner technologies (to produce electricity, for instance), which is equivalent to having technology with a cheap abatement cost. It looks like Europe has followed this path, by investing in green technology and auto-imposing self-constraints on emissions. Following this, I assume that there are two groups of countries: rich countries, with high environmental concern and low cost abatement technology; and poor countries (also called outsiders), with lower environmental concern and expensive abatement technology. Also, I assume that for rich countries it is profitable to form a coalition, where it is not for poor ones.

Second, I add the possibility of transfers between countries, which creates the potential to enlarge the existing coalition in a fashion similar to Carraro and Siniscalco (1993). However, I consider technology transfers (instead of monetary transfers, as Carraro and Siniscalco do), which can change the game itself. The final addition to the baseline model is that I allow the coalition to impose a border tax on non-signatories. As expected, the idea behind this is to deter free-riding (and possibly induce accession), similarly to the intention of the trade ban in Barrett's (1997) paper.

Through this set-up, I want to study if a small initial coalition of green countries can induce the formation of the grand coalition. The idea behind this is the same as the successful mechanism implemented in the Montreal protocol and its subsequent amendments (for a detailed explanation, see Barrett's book (2003)). In this case one country, the United States, was unilaterally willing to cut emissions and consumption of ozone-depleting substances. They were well aware that at first, leakage coming from trade could reduce their effort results, but knew that if other big economies would follow suit, the gains (both nationally and globally) would be much larger. Therefore, they were willing not only to unilaterally ban the use and production of CFCs, but also to ban trade of these substances. This second component, plus the fact that they were also willing to help developing countries to switch to new and clean substances, pushed other countries to join what would become one of the most successful IEAs in history.

Obviously, it would be appealing to do a similar thing with greenhouse gases (GHG). However, while tackling the production and trade of CFCs is a simple task, it is quite impossible for CO_2 emissions, since they are embedded in almost every product we trade. Of course, completely banning trade seems out of the question. The 'sticks' used in Montreal are not credible in an IEA concerning CO_2. To avoid this obstacle, we can use a border tax, imposed on goods coming from non-signatories countries, in order to deter free-riding and induce accession. This will be the case in the present paper, acknowledging that this could bring on a trade war. With respect to this, I assume that coalition members can impose a border tax and non-signatories do not retaliate. I then drop this assumption and analyse how credible this is and which cases it actually holds true for.(2) Furthermore, in a recent paper by Nordhaus (2015), which studies a DICE model with 15 specific regions, the author also assumes that it is possible to impose a border tax without retaliation and shows that this tax induces countries to join the coalition. I find a similar result, plus an upper limit for this tax imposed by the condition of inducing the grand coalition.

With all this in mind, I analyse if a border tax imposed by signatories, a technological transfer, or both can induce the grand coalition. To do so, I study the stability of the grand coalition for different parameter scenarios.(3) I analyse the different cases in which either the tax, the transfer, or both together sustain the grand coalition. These solutions

(2)The results found here are also in line with those of Anouliés (forthcoming), where the author shows that for some parameter configurations there is no retaliation.

(3)In order to study a meaningful set of parameters, I assume that parameters are such that in all possible cases countries are trading with each other. Also, parameters' space is delimited in order to have a stable coalition of rich countries.

depend on, among others things, the countries' environmental damage (or concern). This result fits with the aforementioned paper, since Nordhaus also finds mixed effects. Then I assume that both instruments are needed and I find the optimal recipient group. I give an example of this case, which portrays how both instruments work together.

Finally, I investigate the case in which we can discard the Minimum Participation Clause (MPC).(4) These clauses are legal tools that help IEAs reach the desired equilibrium as in, for example, the Montreal case. However, in this particular case, the MPC was not too high.(5) Concerning the CO_2 problem, although there is a consensus about the damages arising from global warming, there are differences in opinion among countries regarding how these damages will hurt them specifically. Moreover, the damage level is uncertain and eventually comparable to the cost of switching to the required clean technology. Uncertainty about the gains and costs of such an agreement makes the political decision a much harder one to pursue and puts the MPC on a level that might not be politically feasible to reach. Therefore, I perform a theoretical exercise by verifying what would be needed in order to reach the grand coalition without an MPC. The goal is to illustrate the mechanics of what would have to be done to lower the necessary MPC.

The rest of the paper is structured in the following way: Section 2 presents the model and solves the firm maximization program. Section 3 analyses if a border tax, a technological transfer, or both induce the formation of the grand coalition. Section 4 finds the optimal recipient group and analyses the special case with no MPC. Section 5 concludes.

2 The model

I develop a two-stage game built on Barrett (1997). In the first stage, each country chooses whether or not to be part of the coalition of size k, S_k. There are N countries that are asymmetric in two dimensions. First, the technology they have access to can offer a low marginal cost of abatement (σ_L, good technology) or a high one (σ_H, bad technology). Second, countries differ in terms of the marginal damage from emissions they suffer from, noted ω_j, for country j. I will be using marginal damage from emissions and environmental concern as synonyms; in this line, I will also be referring to countries with higher environmental concern as 'greener'. I then define two types of countries.

(4) The Minimum Participation Clause states that the treaty becomes binding when an amount of countries greater than or equal to this minimum have signed the agreement. Since IEAs are faced with multiple equilibria (typically two), in order to induce the optimal one, treaties include an MPC. The objective is to make sure that signatories do not suffer when few countries have signed the agreement.

(5) In order for the Montreal protocol to become binding for parties, it would have to be ratified by at least 11 countries accounting for at least two-thirds of the 1986 level of global consumption of the controlled substances.

The 'rich' are those with the good abatement technology, σ_L, and a high marginal damage from emissions, ω_H. The 'outsiders' or 'poor' have the bad abatement technology, σ_H,and a lower environmental concern than the rich: $\omega_j < \omega_H$, where ω_j, the marginal damage from emissions of country j, is a continuous variable.

Departing from Barrett's (1997) model, I assume that if a country enters the coalition, it will fully abate ($q_j = 1$). If not, it will abate nothing ($q_j = 0$). Hence, for country j, joining the coalition and having $q_j = 1$ become synonyms.(6) In the non-signatory case, this assumption poses no real restriction, since $q_j = 0$ is optimal for these countries (using some standard assumptions), which is not the case for signatories. The assumption can be understood as follows: countries choose between two types of technology – clean and dirty. The clean one does not pollute at all, which translates in this model into full abatement; the dirty technology pollutes, which is equivalent to no abatement.(7) In the present framework, this is equivalent to a σ_L below some threshold and σ_H above some other threshold.(8) Also, this feature has two good implications: First, I discard by construction the case of having coalitions (especially the grand coalition) that do not abate, or who operate quite close to BAU when they do abate, as the literature has shown (as in Barrett (1994) and Eichner and Pethig (2013)). Second, it makes the model more tractable with clear-cut results.(9) On another side, if countries' decisions are binary, it makes the coalition formation a harder process, since becoming a member of the IEA implies full abatement for the joining country. Finally, signatories can not punish a country for leaving the coalition by increasing emissions.

Focusing on abatement technology rather than on emissions is important: From a political point of view, it can more easily lead to an enforceable IEA since the technology being used is more easily verifiable than total emissions. Furthermore, it allows us to consider technology transfers, which change the recipient country's incentives to join the coalition, and therefore the game itself, which is part of the overall plan.

Countries inside the coalition tax imports, at a rate t, of goods produced in non-signatory countries. The main obstacle to reaching a meaningful IEA comes from having carbon leakage, because it provides countries strong incentives to free-ride. The border tax is, then, a credible tool for hindering leakage. In this set-up I assume that only signatories tax goods coming from non-signatories and that the latter do not retaliate with another tax on signatories' goods. First, general results will be derived using this

(6)Conversely, for country j, not joining the coalition and having $q_j = 0$ are also synonyms.

(7)Alternatively, it can assumed that there are two emissions levels – low and high. It is easy to see that this framework is equivalent to the present one, in which I have 'normalized' the low emission level to zero.

(8)These thresholds are calculated explicitly in Section 3.

(9)Binary choices can also be found in the literature, as for example, in Barrett (2003) and Heal (1994).

framework and, in the following subsection, I drop this assumption and analyse if non-signatories prefer to retaliate.

In the second stage, firms move by choosing simultaneously their segmented outputs, all within a Cournot-Nash set-up. It is a perfect information model in the sense that countries perfectly know their costs and gains, as well as those of other countries. Focusing now on the solution of the game at hand, I proceed using backward induction.

2.1 Firms' choices

There are N countries and N firms (one per country) that produce an homogeneous traded good and a transboundary pollution. The inverse demand in each country j is given by $p\left(x^j\right) = 1 - x^j$, where x^j is consumption in country j. Costs of the firm in country j are $C\left(\sigma_j, x_j, q_j\right) = \sigma_j q_j x_j$, where x_j is the total output of the firm, $\sigma_j \in \{\sigma_L, \sigma_H\}$ is its marginal abatement cost, and $q_j \in \{0, 1\}$ is the abatement standard chosen by the government of country j, taken as given for the firm in this country. Emissions by firm j are $x_j\left(1 - q_j\right)$: if abatement is maximal, emissions are zero, and if no abatement is undertaken, emissions are equal to output. The marginal abatement cost of firm j could reflect the technology used – to produce electricity, for example – in country j. Therefore it could be thought as the (accumulative) efforts undertaken by a country to be greener.

Firms choose their output for each market simultaneously. Transport costs are zero, and each firm takes its own abatement standard and the segmented outputs of other firms as given. Firm j chooses a quantity to produce and ship to market i, x_j^i, so as to maximize its profit π_j:

$$\max_{x_j^i \geq 0} \pi_j = \sum_{i=1}^{N} (1 - x^i - t_j^i - \sigma_j\, q_j) x_j^i \tag{2.1}$$

with $t_j^i = t$ if $i \in \mathcal{S}_k \wedge j \notin \mathcal{S}_k$, and $t_j^i = 0$ otherwise. There are N first order conditions for firm j:[10]

$$1 - x^i - t_j^i - \sigma_j\, q_j - x_j^i = 0 \qquad \forall i, j \tag{2.2}$$

Taking into account the fact that $q_j \in \{0, 1\}$ and solving the system of equations formed by the N first order conditions for the N firms, we get:

$$x^i = \begin{cases} \frac{N - \sigma_{\mathcal{S}} - (N-k)\cdot t}{N+1} & \text{if } i \in \mathcal{S}_k \\ \\ \frac{N - \sigma_{\mathcal{S}}}{N+1} & \text{if } i \notin \mathcal{S}_k \end{cases} \tag{2.3}$$

where k is the size of the coalition, and $\sigma_{\mathcal{S}} = \sum_{j \in \mathcal{S}} \sigma_j$ (the sum of the coalition's marginal

[10] I only consider those situations where firms produce positive quantities in equilibrium.

abatement costs). Replacing this result into the FOCs, we have:

$$x_j^i = \begin{cases} \frac{1+\sigma_{\mathcal{S}}-(N+1)\sigma_j+(N-k)\cdot t}{N+1} & \frac{1+\sigma_{\mathcal{S}}-(k+1)\cdot t}{N+1} & \text{if } i \in \mathcal{S}_k \\ \frac{1+\sigma_{\mathcal{S}}-(N+1)\sigma_j}{N+1} & \frac{1+\sigma_{\mathcal{S}}}{N+1} & \text{if } i \notin \mathcal{S}_k \\ \uparrow & \uparrow & \\ \text{if } j \in \mathcal{S}_k & \text{if } j \notin \mathcal{S}_k & \end{cases} \tag{2.4}$$

From where we can calculate the firm profit π_j and production x_j:[11]

$$\pi_j = \begin{cases} \frac{N\cdot\aleph_j^2+k(N-k)t\left(2\aleph_j+(N-k)t\right)}{(N+1)^2} & \text{if } j \in \mathcal{S}_k \\ \frac{N\cdot\aleph_2^2+k(k+1)t\left((k+1)t-2\aleph_2\right)}{(N+1)^2} & \text{if } j \notin \mathcal{S}_k \end{cases} \quad \text{with} \quad \begin{array}{l} \aleph_j = 1+\sigma_{\mathcal{S}}-(N+1)\sigma_j \\ \aleph_2 = 1+\sigma_{\mathcal{S}} \end{array} \tag{2.5}$$

$$x_j = \begin{cases} \frac{N\cdot\aleph_j+k(N-k)t}{N+1} & \text{if } j \in \mathcal{S}_k \\ \frac{N\cdot\aleph_2-k(k+1)t}{N+1} & \text{if } j \notin \mathcal{S}_k \end{cases} \tag{2.6}$$

2.2 Countries' choices

Country j's net benefits are the sum of firm j's profits, plus the surplus of country j's consumers, less the environmental damage suffered, plus border taxes collected, if that is the case. Given demand specifications, the consumer surplus is equal to $(x^j)^2/2$. Pollution is assumed to be a pure public bad, and aggregate emissions are given by $\sum_{i=1}^N x_i(1-q_i)$, where x_i is the total output of firm i. The constant marginal environmental damage is equal to ω_H for the rich countries, and to $\omega_j < \omega_H$ (abusing the notation) for a country j belonging to the group of *outsiders*. If country j is a signatory country (meaning that $q_j = 1$), the border taxes collected are equal to the tax rate t, times the imports from non-signatories (all countries i such that $q_i = 0$). With these, country j's profit is:

$$\Pi_j = \pi_j + (x^j)^2/2 - \omega_j \left[\sum_{i=1}^N x_i(1-q_i)\right] + t\cdot q_j \cdot \sum_{i=1}^N x_i^j(1-q_i) \tag{2.7}$$

where π_j, x^j, x_i and x_j^i are expressed in terms of k and model parameters (given by equations (2.3) and (2.4)). Countries choose whether or not to join the coalition, and therefore whether to fully abate or not at all, by maximizing this profit.

[11] Mathematical development of equations (2.2), (2.3) and (2.4) are in Appendix A.

2.3 Coalition formation

Countries decide either to join the coalition or stay outside of it, in which case they act as singletons. Countries join the coalition if it is profitable for them to do so; they decide this by comparing their profit when belonging to a coalition of size k to their profit when staying outside it (resulting in a coalition of size $(k-1)$). Being profitable can have two meanings. We can either assume that it is profitable for each country, individually and *without* profit sharing among the coalition, or we can assume that the countries belonging a coalition share their profits according to some sharing rule. Since the second case is better for the coalition compared to the individual option, I assume that the coalition forms according to this criteria. Nevertheless, noting that country profit (Eqn. 2.7) depends linearly on ω_j, some analyses are made using an individualistic approach and then extended to the one where signatories share gains.

In order to analyse the coalition stability, I rely on the well-known concepts of Internal Stability (IS) and External Stability (ES)[(12)]. In the case of the grand coalition, we only need to verify for IS, i.e. checking that the sum of profits of coalition members is greater than the sum of their *outside options* (meaning the profit each country would make if it were to leave the coalition of k countries and a coalition of the remaining $(k-1)$ countries formed). The IS condition can be represented as

$$\sum_{j \in \mathcal{S}_k} \Pi_j^s > \sum_{j \in \mathcal{S}_k} \Pi_j^{oo} \tag{2.8}$$

where Π_j^s is the country profit of a signatory (belonging to the coalition of size k: $\mathcal{S}_k$) and Π_j^{oo} is the country outside option (for the same group of countries). If inequality (2.8) is divided by k (the coalition size), we can talk of the mean coalition profit and the mean outside option, which will turn out to be more convenient.

I also assume that the parameters of the model are such that there exists a small coalition of rich countries, which is stable (i.e. IS and ES). The group of rich countries is denoted by $\mathcal{D}$. This coalition can be expanded, especially to try to induce the grand coalition, by the use of a border tax and/or technological transfers. The latter means that some r countries can receive a technological transfer that reduces their marginal abatement cost from σ_H to σ_L. This recipient group is denoted by $\mathcal{R}$. These technological transfers can be thought of as international aid from rich countries to some other countries in order to shift from carbon-intensive power sources toward more eco-friendly ones – for example, to change how electricity is produced. This one-time transfer costs K per recipient.

(12)Internal and External Stability as defined by D'Aspremont et al. (1983) and subsequently used by a substantial literature. IS and ES mean that for a given coalition, no member prefers to leave and no non-signatory wants to join, respectively.

3 Inducing the grand coalition

As stated in the Introduction, I want to verify if using two instruments is better than using one or none in order to induce the grand coalition. The intuition for this goes as follows: technological transfers (TT) lower costs. However, if benefits are low, poor countries will not come in. A border tax (BT) raises the cost of being out, but if costs of abatement are high and environmental concern for these countries is low, poor countries might prefer to be out even when facing the BT. So what could matter is the combination of a BT and TT.

Following this idea, I analyse four situations: Case (1) An IEA with no BT and no TT; Case (2) IEA with BT but no TT; Case (3) IEA with TT but no BT; and Case (4) IEA with BT and TT. With all this, I focus on conditions needed for the grand coalition (GC) to be Internally Stable (IS) in these four cases.

Therefore, the idea is to analyse if the use of one or two instruments enlarges the parameter space within which the grand coalition is stable. I am interested in seeing in which condition using both instruments (Case (4)) works better, and the same idea for Cases (2) and (3). In this context, to 'work better' means that the parameter space in which the GC holds is enlarged. In other words, it means that using one or two instrument makes the GC stable, where this was not the case without the help of these instruments. At the same time, it means that those parameters that already sustained the GC still do, but with greater profits for joining the GC (the coalition gets 'stronger').

As mentioned in section 2.1, parameters are such that the firm's program has interior solutions, which translates into $x_j^i > 0$. On the other hand, and as mentioned in the previous section, I explore those cases (parameter values) where we are in the presence of a (small) initial coalition composed of only rich countries. This means that rich countries are sufficiently green and enjoy cheap technology (or some combination of both). At the same time, this means that outsiders or poor countries care so little about the environment and/or their abatement technology is so expensive that they prefer to stay out of the coalition. Consequently, we have the following conditions:

1. In any scenario, there are interior solutions for the firm's maximization program. Specifically, x_j^i (production in firm j being shipped to country i) has to be greater than zero.
2. Rich countries want to abate when all other rich countries are abating.
3. Poor countries, knowing that d rich are abating, prefer not to abate.

These last two conditions mean that a coalition of d rich countries is Internally Stable (IS) and Externally Stable (ES). Condition 1 translates into various inequalities, depending on which country is exporting to which other country. These inequalities limit parameters σ_L and σ_H as well as pose limits into the border tax level t. Applying these constraints, we get:[13]

$$\sigma_L < \frac{1}{N-d+1} = \sigma_L^{max} \tag{3.1}$$

$$\sigma_H < \frac{1+d\sigma_L}{N} \tag{3.2}$$

$$\sigma_H < \frac{N+1}{N(N-d+1)} \tag{3.3}$$

$$t < \frac{1+d\sigma_L+(N-d-2)\sigma_H}{N-1} \tag{3.4}$$

Conditions 2 and 3 delimit ω_j and ω_H, for given values of σ_L and σ_H:

$$\frac{\omega_H}{\sigma_L} > \frac{N(2N+1)-(N^2(N-2d+2)+d-1/2)\sigma_L)}{N(N+1)(1-(N-2d+1)\sigma_L)} \tag{3.5}$$

$$\frac{\omega_j}{\sigma_H} < \frac{N(2N+1)+(2N^2-1)d\sigma_L-(N^3+1/2)\sigma_H}{N(N+1)(1+d\sigma_L-(N-d-1)\sigma_H)} = \frac{T_0}{\sigma_H} \tag{3.6}$$

I define a threshold level T_0 for the value of ω_j, which is the maximum level of environmental concern that poor countries can have (for given values of σ_L and σ_H). If ω_j were greater than this value, this would imply that this poor country would be willing to join the d rich country coalition, which is ruled out by assumption.

3.1 Grand coalition stability in the four cases

I will verify that each country prefers to join the grand coalition without any profit sharing. A second option, which derives directly from this one, is to assume that the coalition can share the extra profit with all coalition members (condition 2.8). This second condition is just the sum of each country's incentive to join the coalition: $\sum_{j=1}^{N}(\Pi_j^s - \Pi_j^{oo}) > 0$. Since this condition depends on the value of ω_j (of each member), checking for the coali-

[13] Limits for σ_L (Eqn. (3.1)) and t (Eqn. (3.4)) come from $x_j^i > 0$ in combination to the FOC in Eqn. (2.2). This is verified for different cases: rich country exporting to a poor country, rich to rich, etc. Condition (3.2) comes from $\sigma_L < \sigma_H$. Condition (3.3) is just the combination of Condition (3.1) and Condition (3.2). These conditions have to hold for any coalition size. I have assumed d is much smaller than N, in this case $d < N-3$.

tion IS is equivalent to verifying that the average of ω_j meets these same conditions (those to be found).

The objective of this analysis is to check for which parameter set of ω_j, σ_L and σ_H, the different cases ((1) to (4)) work better than others. Since this is an analysis in a volume (restricted by the constraints previously mentioned), I find different conditions for ω_j, for given values of σ_L and σ_H, for all admissible pairs of σ_L and σ_H (and therefore, from now on, I will just talk about the conditions for ω_j, understanding that these are for given σ_L and σ_H). In order to get a better understanding, I will graph these results in the admissible area of (σ_L, σ_H). This area is bounded by: $0 < \sigma_L < \sigma_L^{max}$, $0 < \sigma_H < \frac{1+d\sigma_L}{N}$ (Ineq. (3.2)) and $\sigma_L < \sigma_H$ (the 45° line). In the coming graphs, the area outside the admissible range will be shaded. Finally, we have to remember that the admissible area (σ_L, σ_H) depends also on the values of d and N. For this matter, different situations of d and N will be plotted.

Case (1)

For Case (1), I have no BT and no TT. I call the incentive to join the coalition:

$$^{(1)}\Delta\Pi_j^{N-1} = \Pi_j^s - \Pi_j^{oo} \tag{3.7}$$

where (1) refers to Case 1 (and hence, (2) will refer to Case (2), etc.), j denotes a country with environmental concern ω_j, and $N-1$ refers to this inequality when the country is evaluating whether or not to enter a coalition of size $N-1$. Therefore, Π_j^s is the profit this country has when joining the grand coalition, and Π_j^{oo} is its outside option, meaning the profit when it stays out and the $N-1$ others form a coalition. Using the firm profit, consumer surplus (CS), and environmental damage equations (the first three terms in equation (2.7)), I get:

$$\begin{aligned} ^{(1)}\Delta\Pi_j^{N-1} = &-\frac{N^2\sigma_H}{(N+1)^2}\big(2+2d\sigma_L+(N-2d-2)\sigma_H\big) \\ &-\frac{\sigma_H}{(N+1)^2}\big(N-d\sigma_L-(N-d-1/2)\sigma_H\big)+\frac{\omega_H\cdot N}{(N+1)}\big(1+d\sigma_L+(N-d-1)\sigma_H\big) \end{aligned} \tag{3.8}$$

The first two terms, which are the firms and CS losses, are always negative. In the first case, this is due to the fact that the firm is now facing full abatement, where in the outside option it faced no abatement at all. It is also lowering its production level, since now there is no carbon leakage (from which it was profiting before). In the case of the CS, since the global production goes down (due to the increase in its global costs), consumers also lose. Finally, the last term is always positive, and it is the suppression of the environmental damage. One thing to notice is that this improvement increases with the carbon leakage. This means that incentives to join the GC increase when the carbon leakage increases. Rearranging terms, we have that for the grand coalition to be stable, the following condition must hold:

$$\frac{\omega_j}{\sigma_H} > \frac{N(2N+1) + (2N^2-1)d\sigma_L + (N^2(N-2d-2) - (N-d-1/2))\sigma_H}{N(N+1)(1+d\sigma_L + (N-d-1)\sigma_H)} = \frac{T_1}{\sigma_H} \tag{3.9}$$

In the same fashion as in the previous subsection, I define a new threshold level T_1 for ω_j. This means that if $\omega_j > T_1$, this country will join the GC.

Case (2)

For Case (2), we now have a border tax t and no technology transfer. As before, I find ${}^{(2)}\Delta\Pi_j^{N-1}$. Before going into the equation, note that for the CS there will be no change with respect to Case (1). Looking at equation (2.3), the CS of a non-signatory does not depend on t. On the other hand, when the grand coalition forms, there is no more tax imposed. Hence, the tax does not affect the CS, and therefore, there is no difference from Case (1).

Replacing the original equations as in Case (1) and writing ${}^{(2)}\Delta\Pi_j^{N-1}$ as a function of ${}^{(1)}\Delta\Pi_j^{N-1}$, we get the following expression:

$${}^{(2)}\Delta\Pi_j^{N-1} = {}^{(1)}\Delta\Pi_j^{N-1} + \frac{tN(N-1)}{(N+1)^2}\Big[2\big(1+d\sigma_L + (N-d-1)\sigma_H\big) - tN - \omega_j(N+1)\Big] \tag{3.10}$$

Looking inside the square brackets, we have a negative term with ω_j. This means that the environmental gain that we found in Case (1) has decreased. This is due to the following: Noting that this gain was bigger the bigger the carbon leakage was, and knowing that when a BT is imposed the carbon leakage decreases, we find that the previous environmental gain decreases as well. On the firm side (which is the other part of the expression inside the square brackets), we have a positive effect if the tax is below some threshold, although not imposing new limits, due to condition (3.4). This effect comes from the reduction of carbon leakage. Therefore, the free-rider in Case (2) is gaining less (from carbon leakage) with respect to what he was doing in Case (1).

Thus we find that Case (2) works for a wider set of parameters $(\omega_j, \sigma_L, \sigma_H)$ if the following condition holds:

$$tN + \omega_j(N+1) < 2\big(1+d\sigma_L + (N-d-1)\sigma_H\big) \tag{3.11}$$

Noting that Eqn. (3.11) restricts a linear combination of t and ω_j, we can derive the condition in which Case (2) is better than Case (1) by verifying the limiting case when $t = 0$ and obtaining the following expression:

$$\omega_j < \frac{2\big(1 + d\sigma_L + (N - d - 1)\sigma_H\big)}{(N+1)} = T_2 \tag{3.12}$$

Hence, if this condition holds, we have that Case (2) can sustain the grand coalition for a wider range of parameters. Again I define a new threshold level T_2 for ω_j (for coming analysis). Knowing these values (σ_L, σ_H and ω_j), we can find an 'optimal' value for the border tax t^* that maximizes the gain for the last country accessing the coalition in Case (2). Maximizing expression (3.10) with respect to t gives us:

$$t^* = \frac{2\big(1 + d\sigma_L + (N - d - 1)\sigma_H\big) - \omega_j(N+1)}{2N} \tag{3.13}$$

Case (3)

For Case (3), we have a technology transfer to r recipient countries (meaning that these will have their marginal abatement cost lowered from σ_H to σ_L) and no border tax. As I just did with Case (2), I will represent the IS condition in function of ${}^{(1)}\Delta\Pi_j^{N-1}$. We now get the following expression:

$${}^{(3)}\Delta\Pi_j^{N-1} = {}^{(1)}\Delta\Pi_j^{N-1} + \frac{r(\sigma_H - \sigma_L)}{(N+1)^2}\Big((2N^2 - 1))\sigma_H - N(N+1)\omega_j\Big) \tag{3.14}$$

The intuition of this result goes as follows: Firms' profits will go down (except for the recipients, for whom this is not the case) because of the reduction in σ_S (recall Eqn. (2.5)). The reduction in production level is the same for signatories and non-signatories, but since non-signatories are producing more, they are the ones losing the most (the term is squared in the firm's profit function). Therefore, the impact for the difference in the firms' side is positive. On the consumer surplus side, we have the exact opposite effect. Finally, for the environmental damage, we again get the effect from the previous case due to the reduction of the carbon leakage gained with the decrease of σ_S. These effects are represented inside the last big brackets in the same order stated here. This term is also multiplied by the difference in the marginal abatement costs (hence the technical improvement) and the number of recipient countries r.

As in Case (2) we get again a threshold level for ω_j, which make Case (3) better than Case (1). Equating for the last big bracket, we have:

$$\omega_j < \frac{\sigma_H(2N^2 - 1)}{N(N+1)} = T_3 \tag{3.15}$$

Therefore, if this condition holds, more transfers (bigger values of r) make the GC stronger. But these technological transfers have a cost: K per recipient. Therefore, we have to verify if rich countries are willing to pay for these transfers. At this point we have two options: either the GC is not IS in Case (1) or it is. If it is IS, rich countries do not have any incentive to make transfers, unless they want to make the GC more stable. On the other hand, if the GC is not IS, rich countries are willing to pay for these transfers provided that gains from switching to the GC (from the non cooperative equilibrium) exceed the technological transfer costs. I then proceed to calculate the gain for rich countries, which is the difference between profit in the GC and the one with the original small coalition of d rich countries:

$$\Delta\Pi_d^{d\to N}(r) = \frac{(N-d)\sigma_H - r(\sigma_H-\sigma_L)}{(N+1)^2}\Big(N - (2N(N-d+1))\sigma_L + (N+1/2)(N-d)\sigma_H - r(N+1/2)(\sigma_H-\sigma_L)\Big) + \frac{\omega_H N(N-d)(1+d\sigma_L)}{(N+1)} \tag{3.16}$$

Since this last expression is decreasing in r, rich countries will just transfer to the minimum amount of r recipients in order to make the GC stable – i.e., ${}^{(3)}\Delta\Pi_j^{N-1}(r^*) > 0$. At this point, we have to check if it is profitable for rich countries to do so, since they also have to bear the cost of the technological transfer K. Therefore, r^* will be the minimum natural number for which these two conditions hold:

$$ {}^{(3)}\Delta\Pi_j^{N-1}(r^*) > 0 \tag{3.17}$$

$$\Delta\Pi_d^{d\to N}(r^*) - (K/d)\cdot r^* > 0 \tag{3.18}$$

If there is no natural number r that make these two conditions hold, then $r^* = 0$, meaning that it is not profitable for the rich countries to transfer technology. Either it does not induce the GC, or it is too expensive compared to the gains attained by switching to the grand coalition.

Case (4)

For this last case, both a border tax t and a technology transfer to r recipients are used. In order to make the analysis simpler, let me define the following expression:

$$ {}^{(i)}\Phi_j^{N-1} = {}^{(i)}\Delta\Pi_j^{N-1} - {}^{(1)}\Delta\Pi_j^{N-1} \tag{3.19}$$

which is no more than the 'gains' of Case (i) with respect to the base case, Case (1). What I mean by 'gain' is that if this value is positive, we have a broader range of parameters where the grand coalition will be stable (I have used this concept in the previous cases without giving it an explicit name). Proceeding as before, we get:

$$^{(4)}\Phi_j^{N-1} = ^{(2)}\Phi_j^{N-1} + ^{(3)}\Phi_j^{N-1} - \frac{2r(\sigma_H - \sigma_L)tN(N-1)}{(N+1)^2} \tag{3.20}$$

It is straightforward from this expression that the gains of Case (4) are less than the sum of the gains of Cases (2) and (3). Hence, for Case (4) to be an improvement of either Case (2) and (3), we need the last term to be less than $min(^{(2)}\Phi_j^{N-1}, ^{(3)}\Phi_j^{N-1})$. Having any of these gains of Case (2) or (3) ($^{(2)}\Phi_j^{N-1}$ or $^{(3)}\Phi_j^{N-1}$) be negative will imply that Case (4) is never the best solution (concerning an improvement of the solution space for the GC). For example, if $^{(3)}\Phi_j^{N-1}$ is negative, it is straightforward that $^{(4)}\Phi_j^{N-1} <^{(2)} \Phi_j^{N-1}$, meaning that Case (2) is better than Case (4).

The intuition for this result goes like this: On one side, having a border tax t reduces the carbon leakage which was the 'driving gain' of Case (3) (when making the transfer). Conversely, the technological transfer r affects the value of σ_S (it decreases it in an amount $r(\sigma_H - \sigma_L)$), which is the driving gain that makes Case (2) better than Case (1) (i.e. $^{(2)}\Phi_j^{N-1}$). Therefore, when I use both instruments, I add them, but this combination also diminishes the original gain of each instrument.

Let us assume that both Cases (2) and (3) are better than Case (1), meaning that both $^{(2)}\Phi_j^{N-1}$ and $^{(3)}\Phi_j^{N-1}$ are positive. Then we could have an improvement moving to Case (4) if the right values for r and t were chosen. Using the condition stated before, we have the following:

$$2tN(N-1) < \underbrace{(2N^2-1)\sigma_H - N(N+1)\omega_j}_{\text{Positive if Case (3) is better than Case(1)}} \tag{3.21}$$

This imposes a new ceiling to the tax t. Furthermore, we should now calculate a new optimal tax value t^{**} that maximizes the expression in Eqn. (3.20). With all this, we should compare the following two possibilities:

$$^{(2)}\Phi_j^{N-1}(t^*) \lesseqgtr {}^{(4)}\Phi_j^{N-1}(t^{**}(r)) \tag{3.22}$$

which would tell us which Case, (2) or (4) is the best solution.

The second condition that we have to verify is the following:

$$2r(\sigma_H - \sigma_L) < \underbrace{2\big(1 + d\sigma_L + (N-d-1)\sigma_H\big) - tN - \omega_j(N+1)}_{\text{Positive if Case (2) is better than Case(1)}} \tag{3.23}$$

In the same fashion as before, this condition imposes a ceiling to r, and we should also verify if it is profitable for rich countries to make transfers in this scenario at all (as we did in Case (3)). Assuming that is the case, we should finally compare the following two expressions in order to choose between Case (3) and (4):

$$^{(3)}\Phi_j^{N-1}(r^*) \lesseqgtr {}^{(4)}\Phi_j^{N-1}(t^{**}(r^{**})) \tag{3.24}$$

Finally, depending on which solution is preferred in expressions (3.22) and (3.24), we can derive whether case (2), (3), or (4) is best. Therefore, for Case (4) we do not have a specific threshold value, but instead, if $\omega_j < T_2$ and $\omega_j < T_3$ (meaning Cases (2) and (3) do better than Case (1)), then Case (4) could be the best solution, depending on the result of the analysis just discussed.

Overall analysis

In this subsection I compare different cases and verify which has the biggest gain for the last signatory to join. This could induce the GC when it was not present, or it could strengthen the GC.

In order to analyse these cases, we should examine all different instances for ω_j – i.e., when it is greater or smaller than combinations of these thresholds. This would imply checking 4! possible orderings. Fortunately, we can note that: $T_1 < T_0$, $T_1 < T_2$, $T_3 < T_0$ and $T_3 < T_2$.(14) We also have that $T_1 < T_3$ iff $\sigma_H > (N+1)/(N^3 + 1/2)$. Finally, $T_2 \lesseqgtr T_0$ depending on a more complex condition that varies with all parameters. Using these four inequalities, we have that $\{T_1, T_3\} < \{T_0, T_2\}$, which yields to four possible orderings of the thresholds. However, if $T_3 < T_1$ we have that $\sigma_H < (N+1)/(N^3 + 1/2)$, which in turn implies that $T_0 < T_2$, and therefore, we end up with only three cases:

A: $T_3 < T_1 < T_0 < T_2$

B: $T_1 < T_3 < T_0 < T_2$

C: $T_1 < T_3 < T_2 < T_0$

(14)These proofs are available upon request. For the first case, I assume that $d < N/2$ and $N > 3$. For the second and third cases I suppose that $2(d+1) < N$. All these assumptions are only sufficient. No condition is need for the fourth inequality.

In other words, depending on the values of σ_L and σ_H (and also d and N), we have these four orderings. This means that we can depict these cases in different areas in the (σ_L, σ_H) plane. Consequently, I use a graphical representation of these cases, depending on the values of σ_L and σ_H (for given values of d and N):

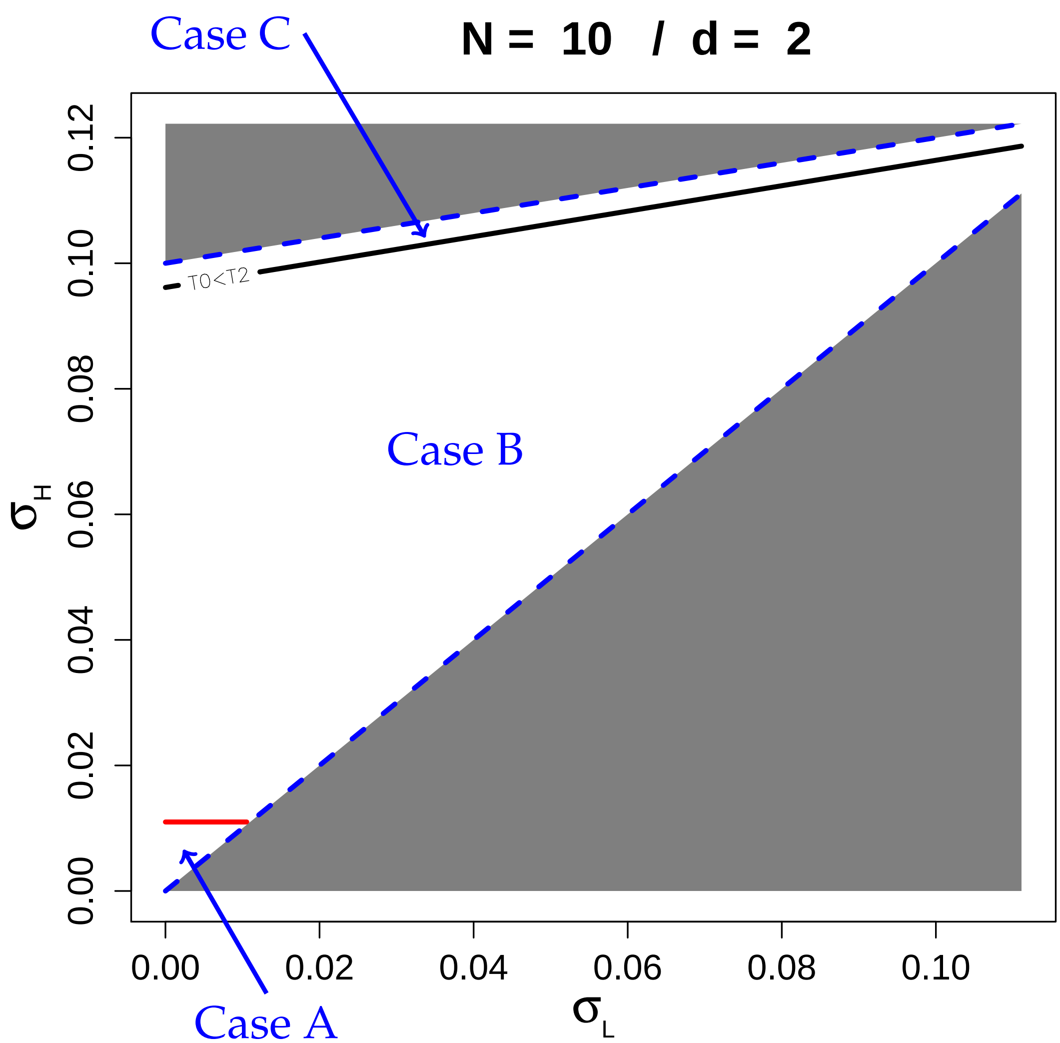
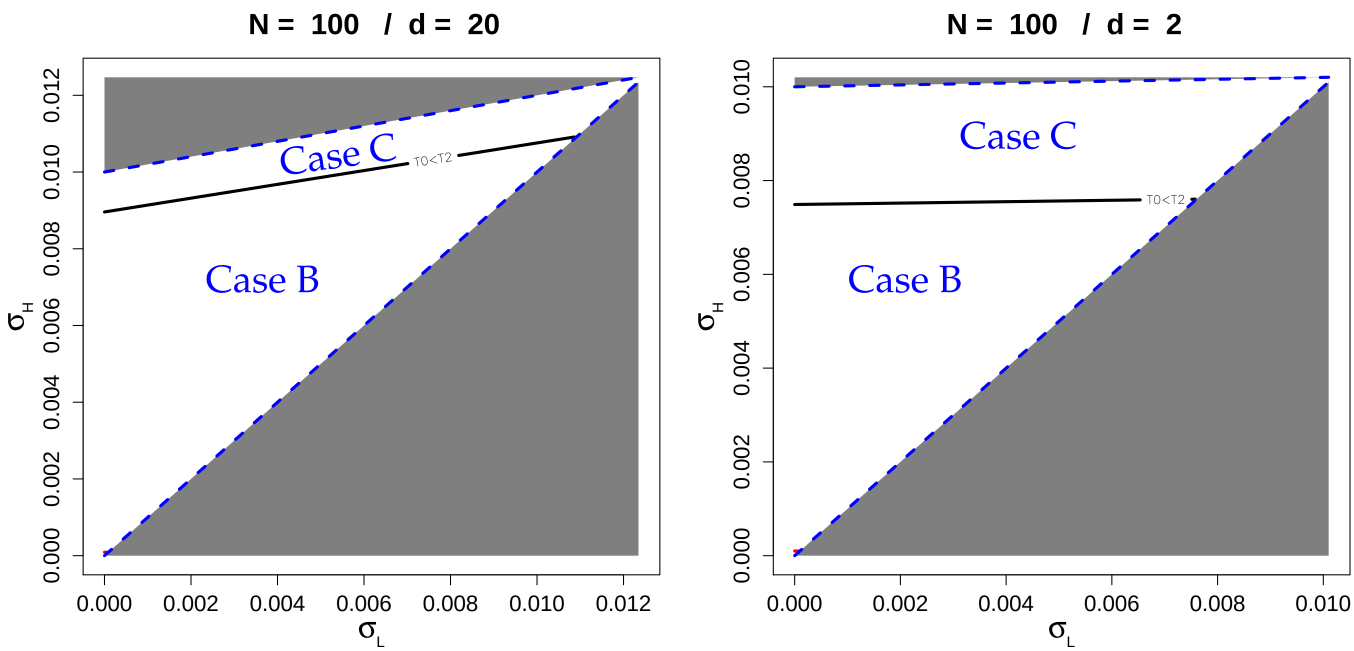


Figure 1: Graphical representation of thresholds' cases.

The dashed lines represent the limits of the admissible values for any pair (σ_L, σ_H). The 45° line simply corresponds to $\sigma_L < \sigma_H$. The upper dashed line comes from condition (3.2). For clarity, the inadmissible area has been greyed. Since $T_1 \lesseqgtr T_3$ only depends on σ_H (and not of σ_L), we get a horizontal line (dividing Cases A and B), which is always below of the $T_0 \lesseqgtr T_2$ line. Furthermore, as N increases the former line goes further south,(15) becoming almost indistinguishable from the corner, as can be noted in the fol-

(15)This comes directly from the result that $T_1 < T_3$ iff $\sigma_H > (N+1)/(N^3 + 1/2)$. Hence, if N grows, the lowing two examples:



(a) (b)

Figure 2: Two more examples of thresholds' cases.

We can also observe that lower values of d (fewer rich countries) make the upper area (Case C) bigger. Now let us focus on the possibilities of ω_j. Following the same order of these three cases, we have (I suggest reading comments from right to left):

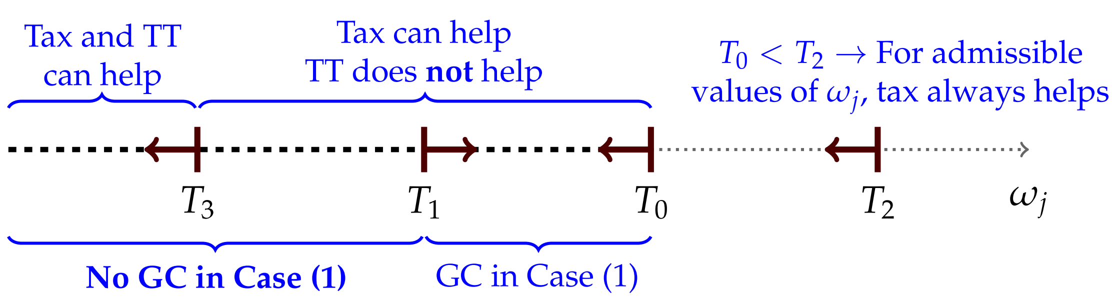


Figure 3: Overall analysis (Case A)

Let us recall that $\omega_j < T_0$ holds by assumption (Inequality (3.6)). To make this point clear graphically, the inadmissible area of ω_j has been represented with a grey dotted line

RHS terms tends to zero rapidly.

(instead of the black dashed line). Observing Cases A and B, when $T_0 \leq T_2$, we see that it is always better to use a border tax in order to expand the range of parameters that sustains the GC (or to make the GC stronger).

Continuing only with case A, where $T_1 < \omega_j < T_0$, we are in the case where the GC exists in Case (1) and, as noted before, a tax can make it stronger. Here, transferring technology does not improve the situation. If $T_3 < \omega_j < T_1$, we find ourselves in the case where there is no stable GC in Case (1) and only the border tax can induce it. Finally, when $\omega_j < T_3$, transferring technology becomes an option (assuming that $r^* > 0$), and the final recipe (BT, TT, or both) will depend on the result of conditions (3.22) and (3.24).

Following the same reasoning for case B, we have:

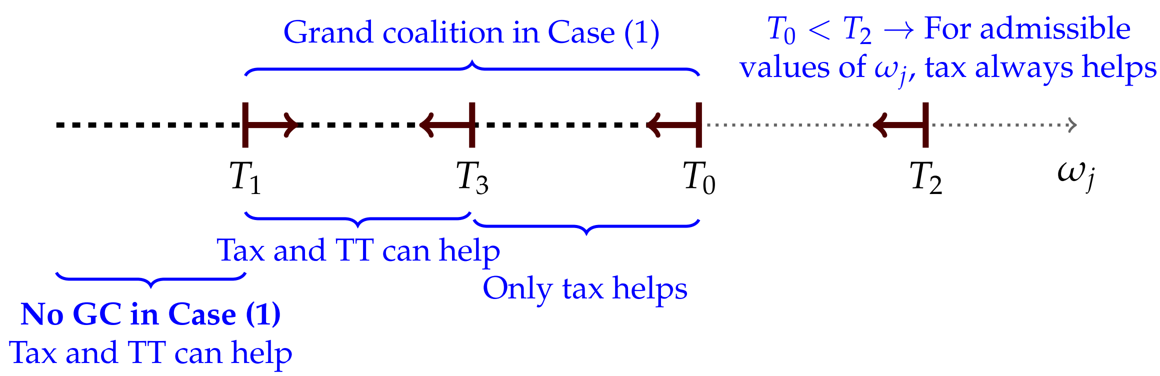


Figure 4: Overall analysis (Case B)

Case B is very similar to case A. The only difference is that we have inverted the order of T_1 and T_3. This translates into adding a range for ω_j where the GC is stable in Case (1) ($T_1 < \omega_j < T_3$) and both a tax and a technological transfer can strengthen it. If $T_3 < \omega_j < T_0$, only the tax can make the GC stronger. Finally, when $\omega_j < T_1$, there is no GC in Case (1), and both instruments can induce it. In third place we have case C, with:

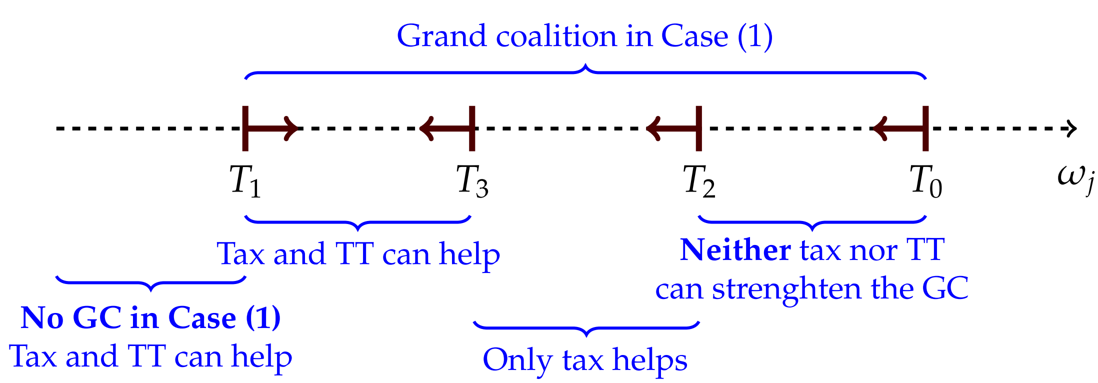


Figure 5: Overall analysis (Case C)

The only difference between case B and C is that in C there is a range of $T_2 < \omega_j < T_0$ where the GC is stable, but neither instrument can improve its stability (this is new compared to Cases A and B). As stated after Figure 2, Case C becomes more relevant when N is big and d is small (Figure 2b). Therefore, when σ_H is close to its upper limit, we are in the presence of this case. In general, though, it seems that Case B is the predominant one.

It is interesting to note that since $T_3 < T_2$, we know that a border tax works in more cases than a technological transfer. In other words, there are more parameter values where the tax works and the transfer does not. When the transfer can work (depending if $r^* > 0$), a tax would work too, the aforementioned Case (4). It is worth recalling that depending on the results of conditions (3.22) and (3.24), it could be the case that the tax is not preferred and the transfer still is.

If the GC is not stable in Case (1), we usually are in the situation where both instruments can work (Case (4)). The only case where this is not true is Case A, when $T_3 < \omega_j < T_1$ and only the border tax helps, but recalling Figures 1 and 2, it seems that Case A is a rare one. The good news is that since $T_1 < T_2$ (as it always is), we know that if there is no GC in Case (1) (i.e. $\omega_j < T_1$), the tax always does better, since $\omega_j < T_2$. The only case where neither instrument can do any better is when $T_2 < \omega_j < T_0$ in Case C, but here we are already in presence of the GC in Case (1).

Up to this point, I have done the analysis assuming countries inside the coalition are not sharing their profits. In other words, for the GC to be stable we need the condition that no country wants to leave, which translates into having the least-green poor country wanting to stay in the GC. This is the value of ω_j discussed so far, which actually is $\min_{j \notin \mathcal{D}, j \notin \mathcal{R}}(\omega_j)$, the least-green poor country that has not received a TT. A second option that can be studied is one where coalition members (in this case, the GC) share their profits. This means that inequality (2.8) holds. Recalling the country profit formulation, we can perform the whole analysis done for far by replacing ω_j with the mean of ω_j just mentioned, getting the same results.

At this point it is worth noting that asymmetry of poor countries can play an important role. If poor countries are symmetric, the no-sharing solution suffices for the general result. On the other hand, if countries are asymmetric, the possibility of sharing profits by signatories changes the result. In the no-sharing case, we need to verify for the least-green non-recipient: It resembles the weakest link problem. But in the sharing case, which is more profitable for the coalition and therefore more interesting, things change. To see how, let us compare two cases: one with symmetric poor countries and one with asymmetric ones, both having the same average ω_j. Therefore, when we are in the presence of TT, non-recipients' environmental concern average $((\sum_{j\notin\mathcal{D},j\notin\mathcal{R}}\omega_j)/(N-d-r))$ will change depending on the recipients chosen. This in an improvement with respect to the symmetric case (since in that case we always get the only value of ω_j) and, as it will be proven in Section 4, the optimal option is to transfer to the least-green poor countries.

3.2 Retaliation tax

So far I have considered that if signatories tax imports coming from non-signatories, the latter will not retaliate with another border tax. In this section I will relax this assumption to examine the incentives of a non-signatory to retaliate. It is good to recall from Brander et al. (1984) that in a set-up of imperfect competition (as this one), countries have the private incentive to impose import tariffs, but they jointly gain more if they enter into a free trade regime. In other words, they face a prisoner's dilemma problem that can be solved with coordination. Therefore, up to this point, in the present set-up, countries have coordinated into free trade.

However, if a country or group of countries deviates from this trade equilibrium and imposes a border tax (even for good motives such as environmental protection), its/their counterpart could retaliate with a similar tax. This is the case that I analyse and, as before, I assume that the coalition will impose a border tax t on products coming from non-signatories and that non-signatories will decide separately whether they will impose a retaliation tax t_2 or not. This means that non-signatories act as singletons and not in a coordinated manner.

Following this idea, a non-signatory country will maximize its own profit with respect to the retaliation tax $t_2 \in [0,t]$. Since this program is solved independently by each non-signatory, which is part of a heterogeneous group, I rename the retaliation tax t_2^j. Recalculating the firm profit, consumer surplus, damages from emissions, and the new tax revenue, and only considering those terms affected by t_2^j, we get:

$$\hat{\Pi}_j(t_2^j) = \frac{(1+\sigma_{\mathcal{S}}+kt_2^j)^2}{(N+1)^2} + \frac{(N-\sigma_{\mathcal{S}}-kt_2^j)^2}{2(N+1)^2} - \frac{\omega_j k(N-k)t_2^j}{N+1} + \frac{\left(k-(N-k+1)\sigma_{\mathcal{S}}-k(N-k+1)t_2^j\right)t_2^j}{N+1} \tag{3.25}$$

where $\hat{\Pi}_j(t_2^j)$ is the part of the non-signatory profit that depends on t_2^j. Note that this expression is independent of t, the border tax chosen by signatories. This is in line with the results of Brander et al. (1984), where optimal taxes of a home and foreign country are independent (they use a two-country model). Verifying the second derivative of this expression with respect to t_2^j, we get that it is always negative. Therefore, making the first derivative equal to zero will give us the optimal tax that a non-signatory j would impose. Doing so, we get:

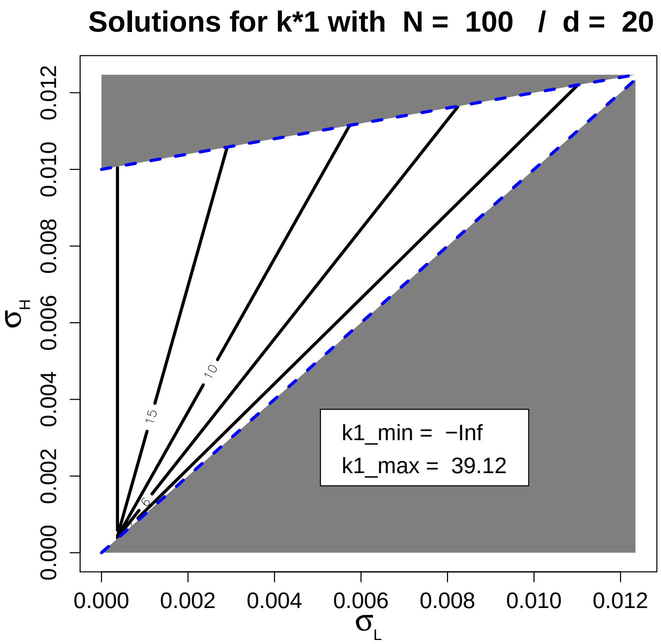
$$t_2^{j*} = \frac{-\Big(3k + (N+1)(N-k+1))\sigma_{\mathcal{S}} - \omega_j(N-k)(N+1)k\Big)}{(3k - 2(N+1)(N-k+1))k} \tag{3.26}$$

Due to the complexity of this expression I will perform two simple analyses. In both I verify when this result is less than or equal to zero, meaning that the non-signatory prefers not to retaliate. First, I check when retaliation is not preferred in the (σ_L, σ_H) plane, as in Figure 1. Note that this expression also depends on the environmental concern ω_j and on the coalition size k (previous analyses were made only for the grand coalition). Realizing that the environmental concern only *helps* to make the retaliation tax equal to zero, I will assume the worse case, meaning that $\omega_j = 0$, and check which values of k this inequality holds for. Since we know that the denominator is negative (it is the second derivative of the original program), we find that this condition translates into the following:

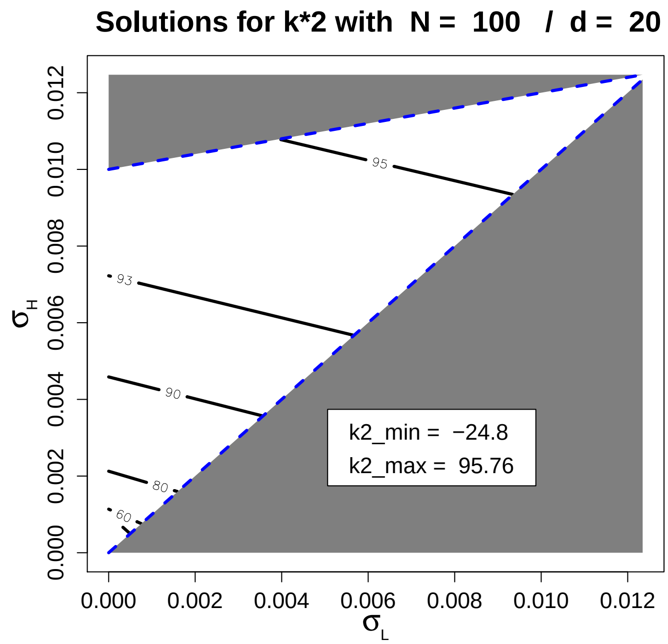
$$k(3 + (N+4)\sigma_{\mathcal{S}}) \leq (N+1)^2\sigma_{\mathcal{S}} \qquad \text{with} \qquad \sigma_{\mathcal{S}} = d\sigma_L + (k-d)\sigma_H \tag{3.27}$$

Solving this quadratic inequality, we finally get k^{*1} and k^{*2}. This means that for $k \in [k^{*1}, k^{*2}]$ the previous inequality holds, and therefore, non-signatories choose not to retaliate.

This area enlarges if $\omega_j > 0$. Since these two expressions are complex again, I plot them in the (σ_L, σ_H) plane, getting:



(a)



(b)

Figure 6: Contours for k^{*1} and k^{*2}

Figure 6a shows the contours for different solutions of k^{*1}. Observing that plot, we realise that these values decrease when the lines 'rotate' clockwise. Take for example $k^{*1} = 15$. For all points to the right of this line – i.e., the pairs (σ_L, σ_H) – we have that $k^{*1} \leq 15$. This means that for coalition of size $k \geq 15$ we know that $k \geq k^{*1}$. In other words, this inequality holds for points in this region. The same reasoning can be performed for k^{*2} in Figure 6b, and we get a second area that for a given k, $k \leq k^{*2}$. In this case, the area begins in one of these contour lines and expands to the northeast. Hence, for a particular coalition size k, we can intersect these two regions, and we get the pairs (σ_L, σ_H), for which $k \in [k^{*1}, k^{*2}]$. In other words, for that region and coalition size k (meaning the triple (k, σ_L, σ_H)), the non-signatory prefers not to retaliate. Note that this result improves if the non-signatory has positive environmental concern (meaning that non-signatories do not retaliate in a larger area). Therefore, for a considerable amount of (σ_L, σ_H) cases, small and medium size coalitions are not retaliated against.

The second analysis focuses on when the grand coalition forms; we want to see what happens if a country leaves the coalition. Would it be in its interest to retaliate? Using the solution in Eqn. (3.26) and knowing that the denominator is always negative, we can again find a threshold for ω_j that tell us when this country is not willing to retaliate. This condition and the according graph in (σ_L, σ_H) are:

$$\omega_j \geq \frac{3(N-1)+(N-5)(d\sigma_L+(N-d-1)\sigma_H)}{N^2-1} = T_4 \tag{3.28}$$

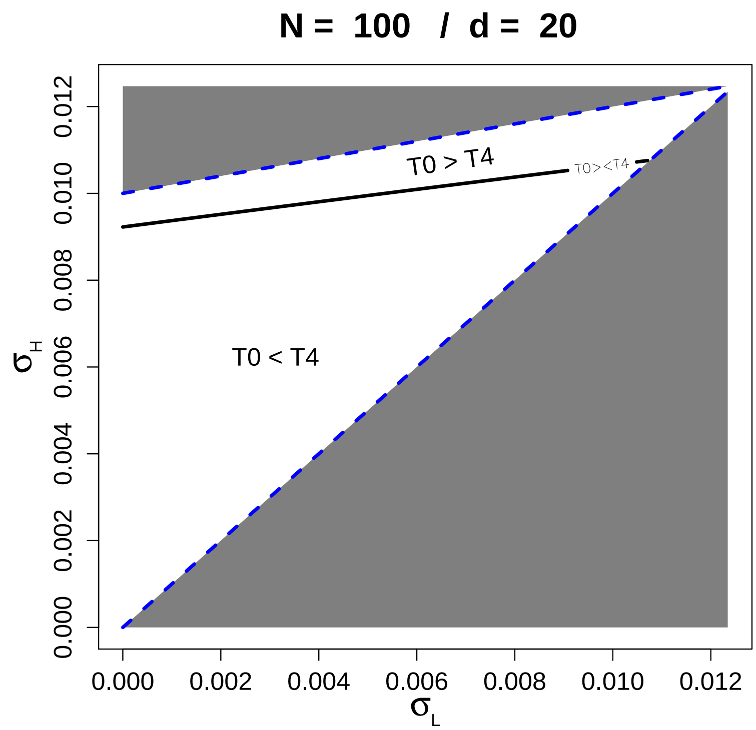


Figure 7: Graphical representation of thresholds T_4.

As we can observe here, this is bad news for the grand coalition. $T_0 < T_4$ means that for that area the non-signatory always prefers to impose a retaliation tax (since $\omega_j < T_0$). Moreover, we also get that $T_1 < T_4$ and $T_3 < T_4$ (for any σ_L, σ_H). The first inequality implies that for those cases where there is no GC in Case(1), and we want to use a border tax to induce it (which now we know always works), non-signatories will prefer to retaliate. In other words, when they do not retaliate ($\omega_j > T_4$) we are always in the presence of the GC in Case (1) ($\omega_j > T_1$). The second inequality is less interesting: It simply means that considering that retaliating is profitable for non-signatories, a TT may or may not help depending on the value of ω_j. For T_2 and T_4, there is no rule: $T_2 \lesseqgtr T_4$. This means that there are cases for ω_j, σ_L and σ_H where a BT and a retaliation tax are optimal, only one of these, or none; all these done by the $(N-1)$ coalition and the non-signatory, respectively.

4 Recipients of technology transfers

In the first part of this section I analyse which country or group of countries should receive a technological transfer if that option would be the optimal situation (hence, we are considering Case (3) or (4)). In the second part, I illustrate the results of using only a border tax or only a technological transfer when compared with the use of both (assuming were are in a situation of Case (4)). The idea here is to give a better understanding of the effects of each instrument and how they work together.

I start by focusing on the country accession. Let us assume that we are in a condition where no country wants to join nor leave: The coalition is stable. If no other country is willing to join the coalition, we consequently know that the greenest outsider is not willing to join. This is simply due to the fact that all outsiders share the same abatement technology σ_H and only differ on their environmental concern ω_j: They can be ordered from the greenest one to the least-green one. Hence, if the greenest one is not willing to join, it is evident from the country profit equation (2.7) that no other country is. It is worth noticing that this is a one-shot set-up, and therefore when I talk about a *sequence* of countries, implying therefore an order, in reality what I am simulating is the thinking process of each country as it decides whether or not it will join a given coalition. We can understand the simultaneous decision process as following: Given a coalition of k countries, the greenest outsider evaluates whether or not it is profitable to join, whatever the other countries do. Obviously, all other countries also think about accessing the coalition at the same time and know that it is profitable for the greenest outsider to join. In that case, it may be profitable for the second-greenest country to join the coalition, no matter what the rest do. In this way, we could have a 'cascading' process ending with all countries joining the coalition. Of course, when I use the word 'cascading', I am not implying a dynamic process, only the strategic reasoning presented. In the same manner, we can clearly see that if this ordered sequence of countries does not produce the 'full cascading' (leading to the grand coalition), no other ordering will. Since all countries know this, they proceed accordingly.[16]

I return to the discussion of the previous section and study the case where both a border tax is implemented and transfers to one or more recipients are possible. The assumption that the BT is implemented is made without a loss of generality (the proof is the same one). No retaliation tax is assumed, though.[17] In other words, Case (4) works the

[16]It could also be considered a dynamic case where countries join sequentially. If this where the case, a discount factor for future gains or losses should be introduced alongside a timing system, which would add more complexity to the model. Since the idea is to keep to model as simple as possible, I do not consider this option, although the reader can visualise this scenario too.

[17]This is assumed for simplicity. Having the possibility of retaliation in the general case escapes the

best. The consequent question that arises is: Which countries should be the recipients? I show that in the presence of a border tax, if a group of $d+r$ (d rich countries and r recipients) produces the cascading, then the set of groups that satisfies this condition always includes the group of outsiders, or poor ones, that are the *least green*. This result can appear counter-intuitive at the beginning, but it should be noted first, that since countries are of the same size, the cost to switch their abatement technology from σ_H to σ_L is always K, regardless of how green the country is. Second, knowing that r countries are recipients and d countries are donors, then $(N-d-r)$ countries are non-recipients, and the goal is to make them produce the cascading. Leaving the greenest countries in the non-recipients group is the best thing to do, since they are more prone to access any given coalition. That means that the countries joining the expanded coalition (donors plus recipients) to form the grand-coalition are the ones that bear the higher abatement costs.

To prove this, I assume that there is a group of r recipients that produce the whole cascading (on top of the d rich countries). Abusing the notation a bit, r will refer both to the amount of recipient countries and to the group of such countries. I show then that if a new recipient group r' is formed, changing one country of the original r group with one less-green country from the non-recipients, then $(d+r')$ also produces the cascading. Iterating the modification of the recipient group, I arrive at $(d+r^*)$, where r^* is the group of least-green outsiders (of the same size of r), which again generates the cascading.

Let us call *Cascading 1* the case where it is supposed that the expanded coalition $(d+r)$ produces the cascading. This means that for each non-recipient i that enters the coalition (in a greenest – less-green 'order' as stated before), the IS condition in (2.8) holds. In the same manner, let us denote *Cascading 2* the case where we have substituted one country from the r group with one country from the non-recipients (this new country being, of course, less green than the replaced one), naming this new recipient group r'. I show that the new expanded coalition $(d+r')$ also produces the cascading, hence:

Proposition 1 *Within the game set-up described above, and with d being the amount of initial rich countries in the coalition and r being a recipient group (of r countries) that produces the whole cascading (Cascading 1) for all i between 1 and $N-d-r$, then the whole cascading is also produced starting from the coalition $(d+r')$, where r' is a less-green recipient group of r countries (Cascading 2):*

$$\underbrace{\Pi_s^{d+r+i} > \Pi_n^{d+r+i-1}}_{\textit{Cascading 1}} \quad \Rightarrow \quad \underbrace{{}'\Pi_s^{d+r+i} > {}'\Pi_n^{d+r+i-1}}_{\textit{Cascading 2}} \qquad \forall i \in \{1, \ldots, N-d-r\} \quad (4.1)$$

where Π_s^{d+r+i} is the non-recipient profit in a coalition of members $(d+r+i)$ (d donors, r re-

scope of this paper and it is addressed as a possible extension.

cipients and i non-recipients), and $\Pi_n^{d+r+i-1}$ is the non-signatory profit of a coalition with one member less (the outside option for the non-recipient i). The prime in $'\Pi$ indicates that we are in Cascading 2, meaning that the recipient group is r' and that the sequence i of non-recipients coming into the coalition has been replaced accordingly. The following diagram shows the non-recipients i sequence for Cascadings 1 and 2:

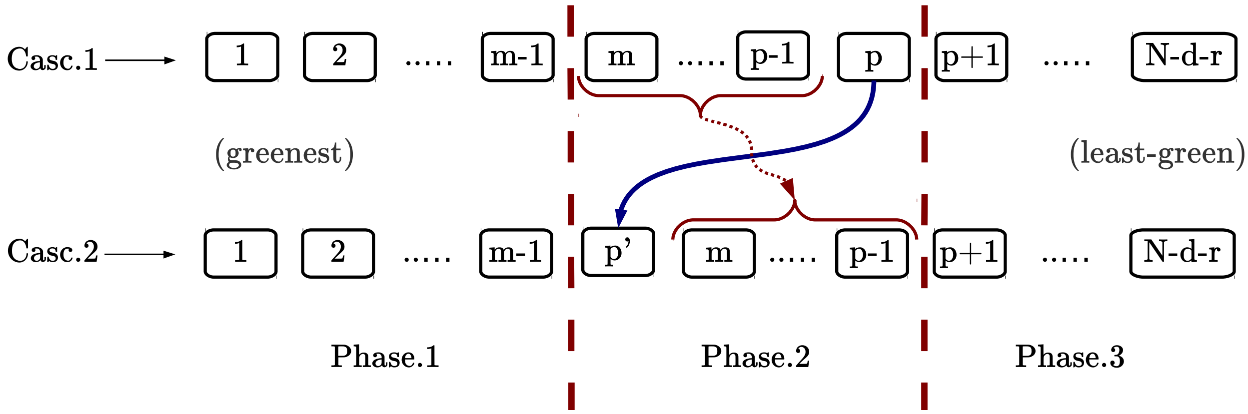


Figure 8: Non-recipients sequence for Cascadings 1 and 2.

As noted in Fig. 8, the non-recipient sequence has been divided into 3 phases. This comes from the fact that we have replaced one country in r, namely [p], which has entered into r', replacing the country $[p']$ that is now in the non-recipient sequence. Note that when a recipient group becomes less green, the corresponding non-recipient group becomes greener, where $[p']$ is greener than $[p]$. Due to this, the sequence has been modified, where phases 1 and 3 are unchanged and the modification only applies to the countries in phase 2. As stated before, $[p']$ is greener than $[p]$ and therefore enters first in the cascading process, as shown in the previous figure.

The proof of the statement in (4.1) consists on showing that for each phase, the following Inequality 1 and 2 hold:

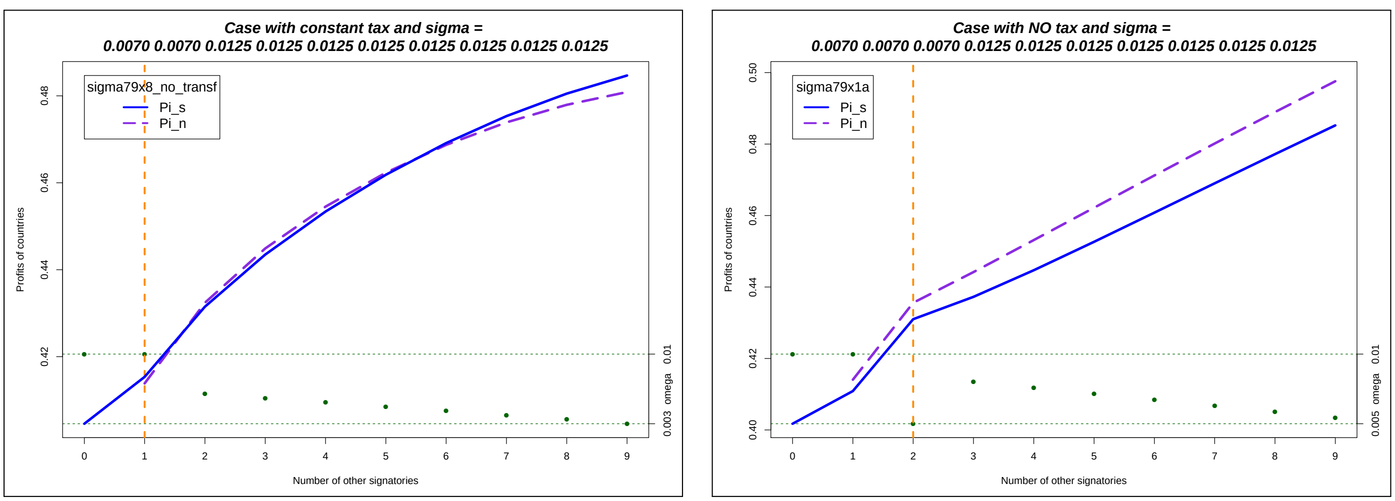
$$\underbrace{\underbrace{'\Pi_s^{d+r+i} \geq \Pi_s^{d+r+i}}_{\text{Inequality 1}} > \overbrace{\Pi_n^{d+r+i-1}}^{\text{Cascading 1}} \underbrace{\geq {}'\Pi_n^{d+r+i-1}}_{\text{Inequality 2}}}_{\text{Cascading 2}} \tag{4.2}$$

A detailed proof can be found in Appendix B. This result states that the set of recipient groups that can produce the cascading always contains the group of the r least-green countries within the outsiders.[18]

[18]Depending on the parameters chosen, this set can contain only one group, r^*, or more than one, but it

The idea behind this proof is the following: by hypothesis we know that Cascading 1 holds. Inequality 1 means that the profit of a country joining the modified coalition (denoted by the prime) is at least as good as the one of a country joining the original coalition. This holds for the first country joining ($i = 1$) and thereafter ($i > 1$), the cascading idea. In both cases the comparison is made using the accession order already mentioned (greener countries first). For Inequality 2 we have the non-signatory counterpart. This just means that the outside option of this country, with the new recipient group r', is worse than or equal to the original case. In other words, with the new recipient group r', joining countries do better (compared to the original situation with r) and their outside options worsen (also compared with the same situation).

In order to illustrate the results of the previous section, I give an example where Case (4) is the best one. I start by showing what happens if one instrument at a time is used, and then when both are used, in the most 'efficient' way (transferring to the least-green countries). To do so, I simulate using 10 countries, in which the initial coalition is composed of the two rich countries. To visualize the decision-making process I use the same type of graphic representation as in Barrett (1997), Carraro (1999), and Diamantoudi and Sartzetakis (2006).[(19)]



(a) Border tax, no transfer.

(b) Transfer, no border tax.

Figure 9: Base case.

Fig. 9a illustrates the case where the border tax is implemented without technological

always includes r^*.

(19) Graphs used in the cited papers considered symmetric countries. Since the goal is to analyse IS (condition 2.8), dividing this inequality by k gives us a *normalized* version of this inequality (on a *per country* basis). Hence, we get: $(\sum_{j \in \mathcal{S}_k} \Pi_j^s)/k > (\sum_{j \in \mathcal{S}_k} \Pi_j^{oo})/k$. The plotted lines are each side of this inequality.

transfers. The border tax rate is assumed to be equal to the marginal cost of abatement outsiders would have incurred if they had abated (σ_H). Fig. 9b presents the reverse case, where one country benefits from the technological transfer ($r = 1$) and no border tax is implemented. Both cases are modelled using the following parameters: $\sigma_L = 0.0070$, $t = \sigma_H = 0.0125$, $\omega_H = 0.100$, $\omega_L = 0.060 \ldots 0.030$.[20]

In both cases, the solid line represents the mean coalition profit and the dashed line the mean outside option. The horizontal axis presents the number of countries in the coalition. Since the countries are heterogeneous, the accession ordering is important. Therefore, to depict this sequence I introduce round dots; these account for ω_j of each country, represented in the secondary y-axis. The d rich countries (equal to two in this example) are the first ones in the graph, on the left. In Fig. 9a no country receives any transfer, and following the previous reasoning (page 26), I only need to test if the greenest – least-green ordering suffices, which is the one depicted. When a transfer is considered, the question lies in the identity of the recipient(s). I plot one option in which the least-green country is the one receiving the transfer, signified by the third dot in Fig. 9b. The rest of the sequence is identical to the one used in Fig. 9a. Choosing another recipient does not change the result. Finally, the value of σ_j for each country (following their order) is printed on the top of the graph, and the vertical dashed line shows the transition from countries with good technology (the expanded coalition) to those without it (non-recipients).

We observe that in the first case, implementing a border tax can alleviate the burden carried by the rich countries by reducing carbon leakage, increasing coalition firms' profits, and getting extra revenues (from taxes), but it needs the use of an MPC in order to trigger the grand coalition formation. The result in the second case is worse, since the absence of the border tax makes the rich countries much worse off and the unique technology transfer does not induce the recipient country to join (and stay) in the coalition.

[20] These values were chosen in order to: a) resemble the examples shown in Barrett (1997), b) have all countries producing a positive amount of good and, c) have a small coalition in the base case.

Let us observe an example of a transfer that induces the GC without the need of an MPC in the following figure.

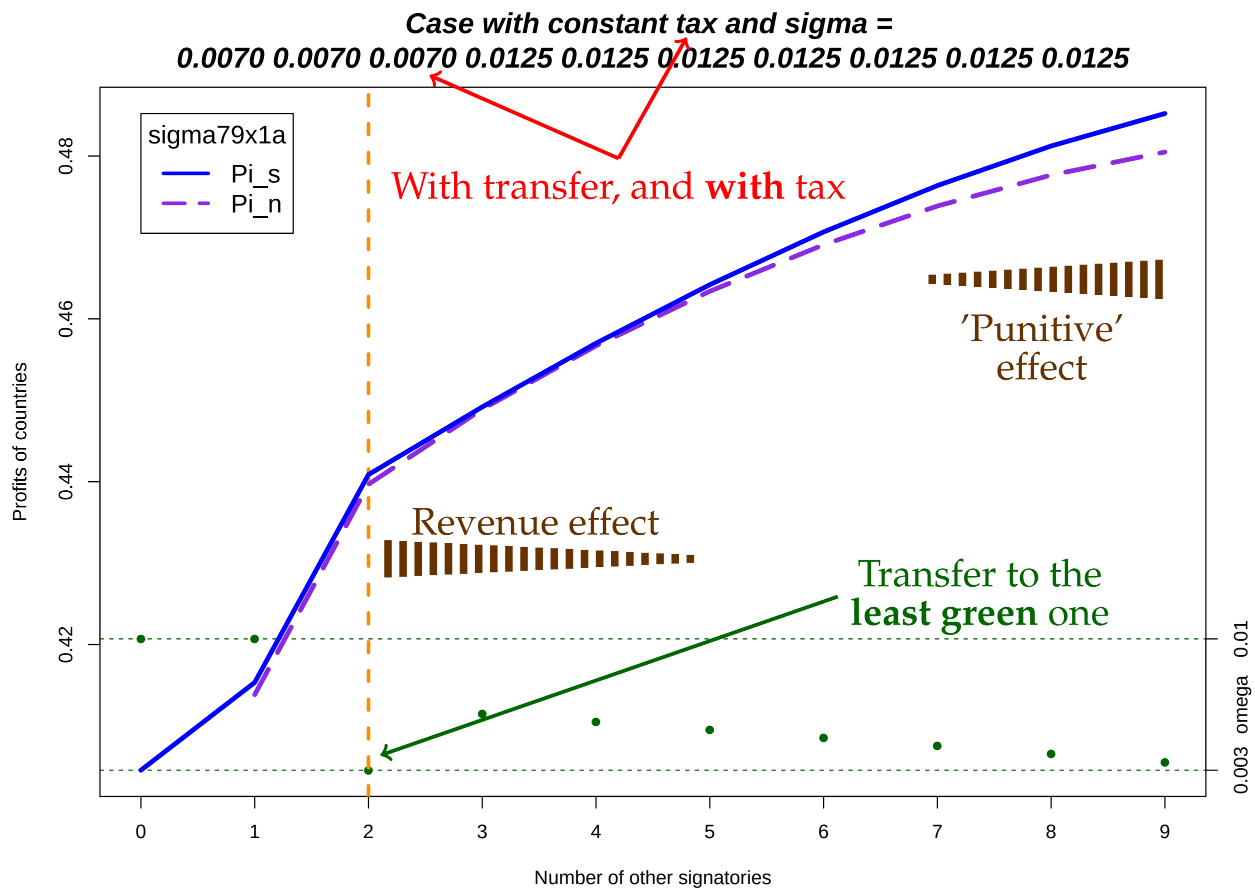


Figure 10: After Tax and Transfers.

I have used the same parameters as in Fig. 9a and Fig. 9b, and the GC without a MPC has been triggered with only one recipient, the least-green outsider. The intuition is as follows: The technology transfer 'buys in' the least-green country, putting it inside the expanded coalition and changing its incentive for abating. The initial coalition is enlarged, and the border tax helps to sustain it. The cascading then occurs, starting with $(d+r)$ countries. In contrast to what happened in Fig. 9a without any transfer, here the IS condition continues to hold even if the revenue effect coming from the border tax diminishes as countries join the coalition. When more countries join the coalition, the border tax has a punitive effect, in the sense that non-signatories' exports are facing a disadvantage with respect to the rest of the world. Nevertheless, after the reduction of the revenue effect and before the apparition of the punitive effect of the border tax, there is a critical period where the IS condition might not hold. This cascading process mimics a mountain crossing[21] (in this case peaking around $k = 3$ or 4), where the outside option is the

[21] As in the Mountain Crossing theorem.

mountain to be crossed and the mean coalition profit is the maximum altitude we reach at each step. With only the border tax, the domino effect stops at some point: The mountain cannot be crossed. But combining the border tax *and* a transfer (when leaving out the greenest countries) allows us to make it through. Both instruments reinforce themselves.

4.1 Disposing of the MPC

Let us assume now that we want to induce the grand coalition without the use of an MPC. This could be due to different reasons – for example, an acceleration of the starting date of meaningful abatement (since we do not need to wait for the MPC amount of countries to ratify the agreement). Therefore, having determined that if the rich countries want to make a technology transfer in order to induce the grand coalition, they have to start with the least-green countries, let us now address the question of how many r^* recipients we need in order to produce the full cascading.(22) Unfortunately, equations developed from the IS condition get much too complex to help create a readable analytical solution for this question. I therefore rely on numerical simulations. I define a $\Delta\Pi^s_{oo}(d, r, i)$ function, which is just inequality (2.8) with all terms put on the left side. Therefore, the IS condition holds if the function is positive. Hence, we have:

$$\Delta\Pi^s_{oo}(d, r, i) = \sum_{j \in \mathcal{S}_k} \Pi^s_j \; - \sum_{j \in \mathcal{S}_k} \Pi^{oo}_j \tag{4.3}$$

Therefore, if for a given value of r, $\Delta\Pi^s_{oo}(d, r, i)$ is strictly positive for all possible i's (d is given), then we get the full cascading. Defining this function makes use of the fact that outsider countries that should be targeted to receive a technology transfer are perfectly known from proposition 1 for each level of r. In Fig. 11 we can observe a continuous version of this function(23) using the same parameters as in the previous examples.

The figure shows the solution area where $\Delta\Pi^s_{oo}(d, r, i) > 0$, and more importantly, its limit $\Delta\Pi^s_{oo}(d, r, i) = 0$. Therefore, it is easy to see that the solution for this case is $r^* = 1$. With $r = 0$ we have the base case where there is no transfer and the border tax is implemented. Following the vertical dashed line at $r = 0$, we can observe that we have the same result as before. For values of i between 0 and 0.8 and then from 3.2 until the end, the value of $\Delta\Pi^s_{oo} > 0$. In the range left in between, it is negative, meaning that for this case we are in need of an MPC if we want to reach the grand coalition (we are not able to cross the mountain). In the case of $r = 1$, we can clearly see that $\Delta\Pi^s_{oo} > 0$ for the full range of i, meaning that full cascading occurs. Following this reasoning, r^* can be found by check-

(22) Of course, it is assumed that this r^* is a profitable solution, as discussed in subsection 3.1, page 11.

(23) This is coming directly from the country profit function. The only 'trick' was that I had to create a continuous version of $\sum \omega_j$ to be used in the damage part of this function. The domain of $\Delta\Pi^s_{oo}$ is restricted to octant I (+++) and with $(d+r+i) \leq N$, which is the area of interest.

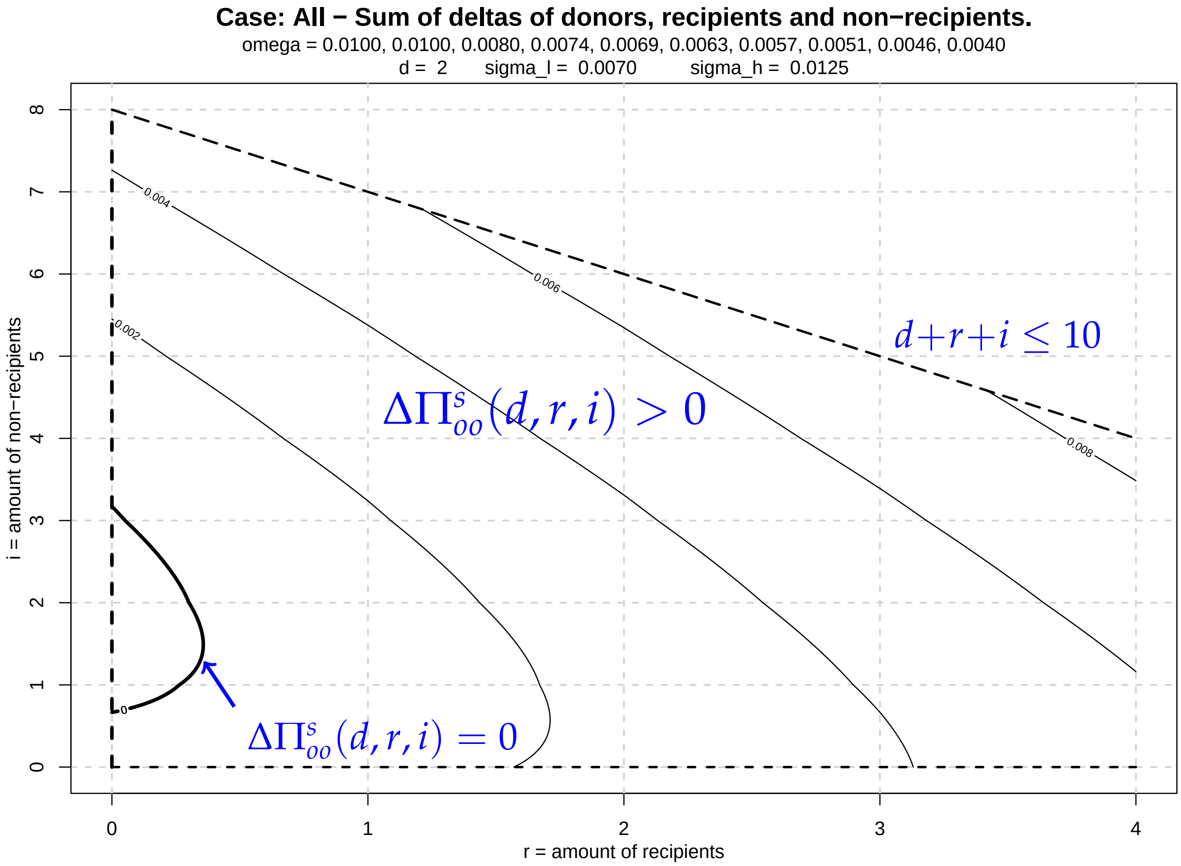
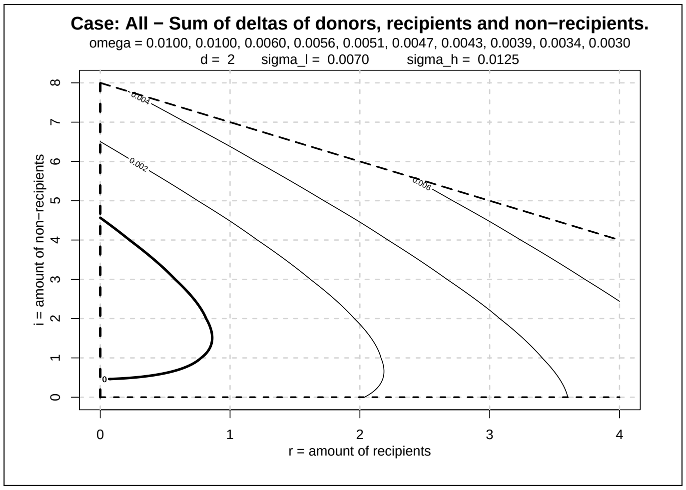


Figure 11: Contour of function $\Delta\Pi^s_{oo}(d,r,i)$, with $d = 2$.

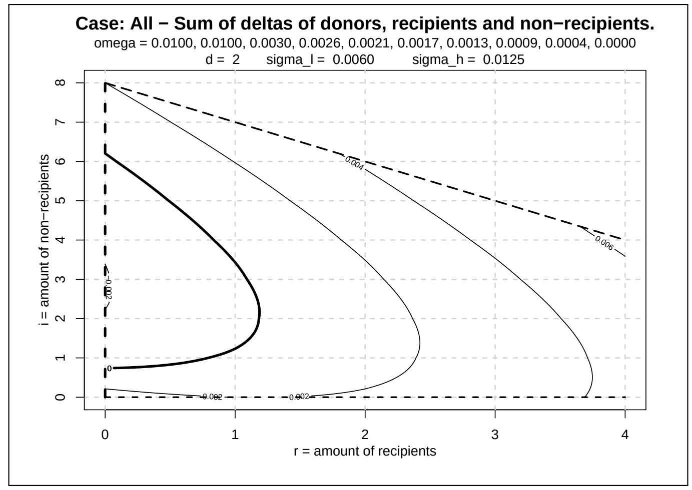
ing when $\partial r/\partial i = 0$ in the implicit function between r and i, given by $\Delta\Pi^s_{oo} = 0$. Hence, we find a maximum $\hat{r}$ (of r with respect to i in this implicit function), and r^* is just the integer solution for $r^* > \hat{r}$. Actually, $\hat{r}$ can cross some integer, say r_1, and r^* can still be equal to $r_1 - 1$. This is due to the fact that we only have to check that, for a given integer level r, $\Delta\Pi^s_{oo}(d,r,i) > 0$ for all possible cases of i integers; this means we only have to check for the points on the grid formed of integer coordinates. Nevertheless, having an analytical solution for $\hat{r}$ would be helpful in order to perform sensibility analysis on r^*, which could be done in future research.

A convenient feature of this graphical representation is that it allows us to analyse what happens if we change some parameters – for example, the environmental concern ω_j of outsiders or the technological levels σ_L and σ_H. Some examples of this are depicted in Figs. 12a–d, where some expected features can be observed.

The lower the environmental concern of outsiders (lower values of ω_j), the harder it is to cross the mountain, meaning that the locus at which $\Delta\Pi^s_{oo} = 0$ switches to the right, and

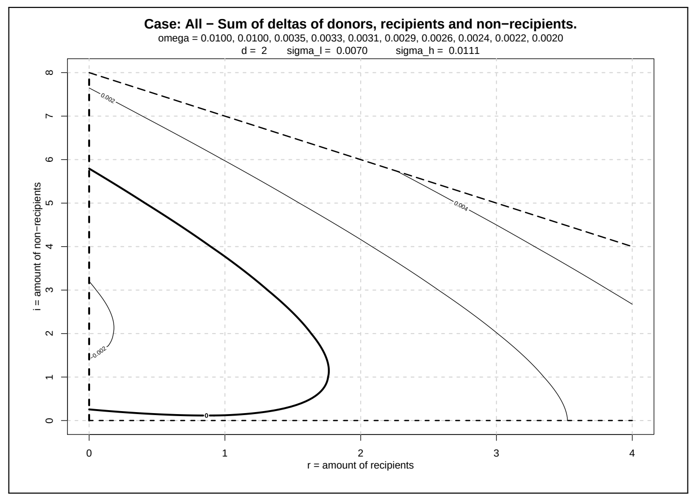


(a) Lower outsiders environmental concern.



(b) Same as (a), but even lower, with better good tech.

(c) Same as (b), but even lower environmental concern.



(d) More even σ_j with cheaper bad tech.

Figure 12: Some examples changing ω_j, σ_L and σ_H.

therefore $r*$ might increase. A similar effect occurs when making abatement technologies more expensive.

5 Conclusions

The present work explores a new approach for reaching an International Environmental Agreement (IEA) by combining two instruments, namely a technological transfer and a border tax. The idea is that the first instrument 'buys in' some countries to the initial coalition through technology transfers, following Heal and Kunreuther's (2011) idea that a set of countries could tip the rest to get to a clean equilibrium. The second one makes sure that these countries could impose costs such as border taxes on non-joiners in order to persuade the rest of the countries to join the coalition, in a way similar to what happened with the Montreal Protocol and its subsequent amendments. Used alone, it might be the case that the border tax is not sufficient to induce the grand coalition, since it needs some critical mass to work. On the other hand, in some situations the transfer cannot work alone either, since it does not deal with the free-rider incentives. The results show that they may exhibit mixed effects when combined. They add to one another, but at the same time they erode their original persuading force. It turns out that depending on parameter conditions, the tax itself, the transfer itself, or both can induce the grand coalition.

I also analyse if non-signatories are willing to retaliate with a similar tax. I solve for which coalition sizes and parameter conditions non-members retaliate. I show that for small- and medium-sized coalitions, the retaliation is not worthwhile; for bigger coalitions it depends on specific conditions of possible deviators. In a recent paper by Nordhaus (2015), he explores, using a DICE model with 15 regions, the effectiveness of a border tax for inducing accession. However, he does not analyse the possibility of a retaliation tax. Hence, it could be explored, using a similar model, if non-signatories would be willing to retaliate. Taking all this into consideration, it could be the case that a big coalition, but not the grand coalition, becomes a stable solution. It is also worth noting that his model is applied to a specific set of parameters (the world divided into 15 regions) and the general result of the present work could help to study how sensitive his results are to changes in this configuration. On the other hand, it would be worth studying the feasibility of technology transfers, combined or not with a border tax, in a model like that.

Another interesting result is that in order to minimize transfers (if they are present in the optimal recipe) and improve the chances of reaching the grand coalition, the recipient countries are those with the least environmental marginal damage, also referred as to the least-green ones. Although it might be a coincidence, it looks like some relation exists with what happened with the bilateral agreement between the United States and China reached last November. Of course in this case, other considerations are at hand, such as the size and emissions levels of these two specific countries.

Finally, a further option can be suggested to expand this work. The strategic implications of being a recipient or not could be analysed. This comes from the fact that non-recipients do not receive any incentive (since they will accede due to the trade pressure). Therefore, this might induce them to join the agreement at an earlier stage and thus get the transfer. This strategic interplay can again be anticipated by the promoters of the IEA and by the rest of the players, enlarging the game set-up and eventually changing or reasserting the previous result.

References

Lisa Anouliés. The strategic and effective dimensions of the border tax adjustment. *Journal of Public Economic Theory*, forthcoming. ISSN 1467-9779.

Scott Barrett. Self-enforcing international environmental agreements. *Oxford Economic Papers*, 46:pp. 878–894, 1994. ISSN 00307653.

Scott Barrett. The strategy of trade sanctions in international environmental agreements. *Resource and Energy Economics*, 19(4):345–361, November 1997.

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

Scott Barrett. *Environment and Statecraft: The Strategy of Environmental Treaty-Making: The Strategy of Environmental Treaty-Making*. Oxford University Press, 2003.

Scott Barrett. Climate treaties and "breakthrough" technologies. *American Economic Review*, 96(2):22–25, 2006.

James A Brander, Barbara J Spencer, and Tariff Protection. Tariff protection and imperfect competition. In Henryk Kierzkowski, editor, *Monopolistic Competition and Product Differentiation and International Trade*, pages 194–206. Oxford Economic Press, 1984.

Carlo Carraro. The structure of international environmental agreements. In Carlo Carraro, editor, *International Environmental Agreements on Climate Change*, volume 13 of *Fondazione Eni Enrico Mattei (Feem) Series on Economics, Energy and Environment*, pages 9–25. Springer Netherlands, 1999. ISBN 978-90-481-5155-4.

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

Lorenzo Cerda. Moving to greener societies: Moral motivation and green behaviour. Working paper, Paris School of Economics - Université Paris 1 Panthéon - Sorbonne, March 2015.

Claude D'Aspremont, Alexis Jacquemin, Jean Jaskold Gabszewicz, and John A. Weymark. On the stability of collusive price leadership. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 16(1):pp. 17–25, 1983. ISSN 00084085.

Effrosyni Diamantoudi and Eftichios S Sartzetakis. Stable international environmental agreements: An analytical approach. *Journal of public economic theory*, 8(2):247–263, 2006.

Thomas Eichner and Rüdiger Pethig. Self-enforcing environmental agreements and international trade. *Journal of Public Economics*, 102(0):37 – 50, 2013. ISSN 0047-2727.

Geoffrey Heal. Formation of international environmental agreements. In Carlo Carraro, editor, *Trade, Innovation, Environment*, volume 2 of *Fondazione Eni Enrico Mattei (FEEM) Series on Economics, Energy and Environment*, pages 301–322. Springer Netherlands, 1994. ISBN 978-94-010-4409-7.

Geoffrey Heal and Howard Kunreuther. Tipping climate negotiations. Technical report, National Bureau of Economic Research, 2011.

Michael Hoel and Kerstin Schneider. Incentives to participate in an international environmental agreement. *Environmental and Resource Economics*, 9(2):153–170, 1997. ISSN 0924-6460.

William Nordhaus. Climate clubs: Overcoming free-riding in international climate policy. *American Economic Review*, 105(4):1339–70, 2015.

SJ Rubio and A Ulph. Self-enforcing agreements and international trade in greenhouse emission rights. *Oxford Economic Papers*, 58:233–263, 2006.

A Firm maximization problem

The firm's profit function to be maximized is the following:

$$\pi_j = \sum_{i=1}^{N} (1 - x^i - t_j^i - \sigma_j q_j) x_j^i \tag{A.1}$$

with $t_j^i = t$ if $i \in \mathcal{S}_k \wedge j \notin \mathcal{S}_k$, and $t_j^i = 0$ otherwise. I only consider situations where firms produce positive quantities in equilibrium. We can get the first order conditions, which are:

$$1 - x^i - t_j^i - \sigma_j q_j - x_j^i = 0 \qquad \forall i, j \tag{A.2}$$

Summing over i, conditions in equations A.2 can be rewritten as:

$$N(1 - \sigma_j q_j) - 2x_j - x_{-j} - t_j = 0 \qquad \forall j \tag{A.3}$$

where $t_j = \sum_{i=1}^{N} t_j^i$ is the sum of the tax rates 'paid' by a product produced by country j, $x_j = \sum_{i=1}^{N} x_j^i$ is the total output of firm j, and $x_{-j} = \sum_{k \neq j} x_k^i$ is the rest-of-the-world output. This can be re-written in a matrix form, getting:

$$\underbrace{\begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{bmatrix}}_{C} \cdot \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}}_{\overrightarrow{x_j}} = \underbrace{\begin{bmatrix} N - N\sigma_1 q_1 \\ N - N\sigma_2 q_2 \\ \vdots \\ N - N\sigma_N q_N \end{bmatrix} - \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_N \end{bmatrix}}_{D} \tag{A.4}$$

Using the Sherman-Morrison formula, we can get $\overrightarrow{x_j} = C^{-1} \cdot D$, being

$$C^{-1} = I - \frac{B}{N+1}$$

where I is the identity matrix and B (of dimension N by N) is a matrix of ones. This gives the general solution of:

$$\overrightarrow{x_j} = \frac{1}{N+1} \begin{bmatrix} N & -1 & \cdots & -1 \\ -1 & N & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & N \end{bmatrix} \begin{bmatrix} N - N\sigma_1 q_1 - t_1 \\ N - N\sigma_2 q_2 - t_2 \\ \vdots \\ N - N\sigma_N q_N - t_N \end{bmatrix}$$

Or equivalently:

$$x_j = \frac{1}{N+1} \left[N\Big(1 - N\sigma_j q_j - t_j + \sum_{i \neq j} \sigma_i q_i\Big) + \sum_{i \neq j} t_i \right] \tag{A.5}$$

Finally, for the country consumption levels x^i, we can use the FOC (eqn. A.2) and we get:

$$x^i = \frac{N - \sum_{j=1}^{N} \sigma_j \, q_j - t^i}{N+1} \tag{A.6}$$

with $t^i = \sum_{j=1}^{N} t_j^i$ being the sum of tax rates *imposed* by country i. Given that only signatories tax non-signatories' products, we have that:

$$t^i = \begin{cases} (N-k) \cdot t & \text{if } i \in \mathcal{S}_k \\ 0 & \text{if } i \notin \mathcal{S}_k \end{cases}$$

which yields to

$$x^i = \begin{cases} \frac{N-\sigma_{\mathcal{S}}-(N-k)\cdot t}{N+1} & \text{if } i \in \mathcal{S}_k \\ \frac{N-\sigma_{\mathcal{S}}}{N+1} & \text{if } i \notin \mathcal{S}_k \end{cases} \tag{A.7}$$

where $\sigma_{\mathcal{S}} = \sum_{j \in \mathcal{S}} \sigma_j$ (the sum of the coalition's marginal abatement costs). Replacing this result into the FOC (Eqn. (A.2)), we finally have the segmented production of firm j shipped to country i:

$$x_j^i = \begin{cases} \frac{1+\sigma_{\mathcal{S}}-(N+1)\sigma_j+(N-k)\cdot t}{N+1} & \frac{1+\sigma_{\mathcal{S}}-(k+1)\cdot t}{N+1} & \text{if } i \in \mathcal{S}_k \\ \frac{1+\sigma_{\mathcal{S}}-(N+1)\sigma_j}{N+1} & \frac{1+\sigma_{\mathcal{S}}}{N+1} & \text{if } i \notin \mathcal{S}_k \\ \uparrow & \uparrow & \\ \text{if } j \in \mathcal{S}_k & \text{if } j \notin \mathcal{S}_k & \end{cases} \tag{A.8}$$

Finally, having x_j, x^i and x_j^i and recalling the equations for the firm profit (π_j), consumer surplus (CS_j), environmental damage (DAM_j), and taxes collected (TAX_j), we get:

$$\pi_j = \begin{cases} \frac{N\cdot\aleph_j^2+k(N-k)t\left(2\aleph_j+(N-k)t\right)}{(N+1)^2} & \text{if } j \in \mathcal{S}_k \\ \frac{N\cdot\aleph_2^2+k(k+1)t\left((k+1)t-2\aleph_2\right)}{(N+1)^2} & \text{if } j \notin \mathcal{S}_k \end{cases} \quad \text{with} \quad \begin{aligned} \aleph_j &= 1+\sigma_{\mathcal{S}}-(N+1)\sigma_j \\ \aleph_2 &= 1+\sigma_{\mathcal{S}} \end{aligned} \tag{A.9}$$

$$CS_j = \begin{cases} \frac{\left(N-\sigma_{\mathcal{S}}-(N-k)t\right)^2}{2(N+1)^2} & \text{if } j \in \mathcal{S}_k \\ \frac{\left(N-\sigma_{\mathcal{S}}\right)^2}{2(N+1)^2} & \text{if } j \notin \mathcal{S}_k \end{cases} \tag{A.10}$$

$$DAM_j = -\omega_j \cdot \frac{(N-k)}{N+1} \cdot \Big((1+\sigma_{\mathcal{S}})N - k(k+1)t\Big) \tag{A.11}$$

$$TAX_j = t \cdot \frac{(N-k)}{N+1} \cdot \Big(1+\sigma_{\mathcal{S}} - (k+1)t\Big) \tag{A.12}$$

B Proof of r^* being the least-green recipient group

As stated in section 4, the proof consists of showing that if conditions for *Cascading 1* hold, then conditions for *Cascading 2* hold too. In order to do so, I will use to following inequalities:

$$\underbrace{\underbrace{{}'\Pi_s^{d+r+i} \geq \Pi_s^{d+r+i}}_{\text{Inequality 1}} \overbrace{\;>\;}^{\text{Cascading 1}} \underbrace{\Pi_n^{d+r+i-1} \geq {}'\Pi_n^{d+r+i-1}}_{\text{Inequality 2}}}_{\text{Cascading 2}} \tag{B.1}$$

Hence, the proof shows that *Inequality 1* and *Inequality 2* hold. The first one says that the profit of an i^{th} country coming into the coalition ${}'\Pi_s^{d+r+i}$ (Cascading 2) is greater or equal than the same profit in the case of *Cascading 1*, where r' has been replaced by the original r and the i^{th} country entering the coalition might or might not be the same as *Cascading 2*, according to the following diagram:

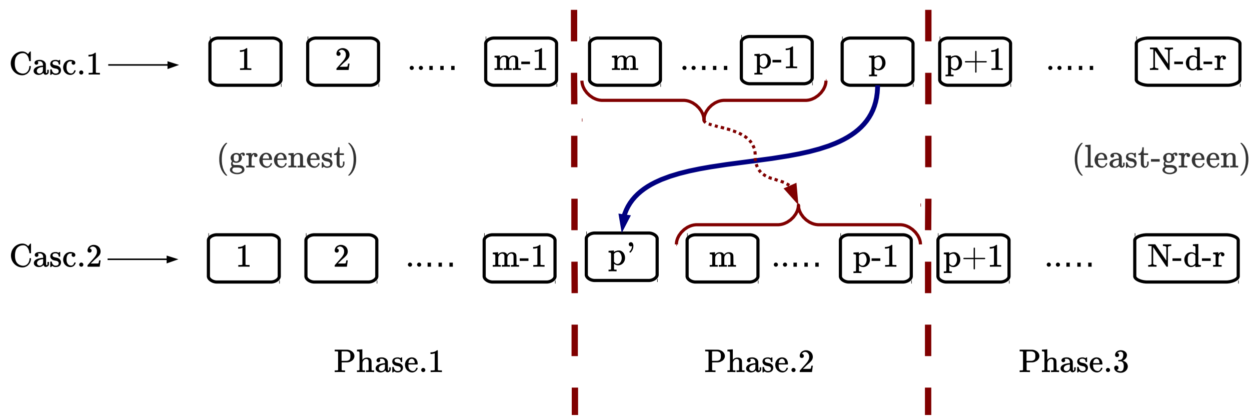


Figure 13: Non-recipients sequence for cascadings 1 and 2.

The same has to be done for *Inequality 2*, where now we have to show that the outside option of an i^{th} country coming into the coalition $\Pi_n^{d+r+i-1}$ (Cascading 1) is greater or equal than its counterpart in Cascading 2. To do so, I will analyse the possible changes in firm profits, consumer surplus, damages, and taxes collected. As shown in the previous picture, I will divide the analysis into three phases: 1, 2 and 3.

First, let us note that the amount $\sigma_{\mathcal{S}} = \sum_{j \in \mathcal{S}} \sigma_j$ (the sum of the coalition's marginal abatement costs) does not change between *Cascading* 1 and 2. This is due to the simple fact that the group of countries with good technology is always of size $(d + r)$. In the same way, $\aleph_j$ and $\aleph_2$ stay constant between those two cases. Therefore, the only difference that may arise comes from the replacement of ω_j, either of countries accessing the coalition (the i's) or those in the recipient group (r or r').

Inspecting the firm profit equation (A.9), it is clear that its value does not change between *Cascading* 1 and 2. The same holds for the consumer surplus and for the taxes collected. Hence, we only have to focus on the damages coming from emissions. Studying this equation shows us that emissions are also invariant between these two cases, since they only depend on σ_S and k. Hence, the only changes comes directly from the term $-\omega_j$ (in Eqn. (A.11)).

Define $\Delta D_1 = {}'\mathrm{DAM}_j^{d+r+i} - \mathrm{DAM}_j^{d+r+i}$, which is the difference in damages of country j between *Cascading* 2 and 1 cases, in the presence of a coalition of size $(d+r+i)$. In the same manner, define $\Delta D_2 = {}'\mathrm{DAM}_j^{d+r+i-1} - \mathrm{DAM}_j^{d+r+i-1}$, which is just the same, with a coalition of size $(d+r+i-1)$. Finally, denote ω_p the corresponding marginal damage for the country p in r being replaced by a less-green country q in r'. And let ω_q be the marginal damage of the replacing country q.

Let us start with the signatories (Inequality 1). In phase 1, $\Delta D_1 = (\omega_p - \omega_q) \cdot \textit{emissions}$, which is positive. In phase 2 we will have a similar case, where the pair of ω's at stake will be: ω_p v/s ω_m, ω_m v/s ω_{m+1} $\ldots$ ω_{m+j-1} v/s ω_{m+j} (recall Fig. 13) and therefore resulting again in $\Delta D_1 > 0$. For phase 3, since the coming countries i's and the group already in the coalition $(d+r+i-1)$ have the same ω's, $\Delta D_1 = 0$. Therefore, for phases 1, 2 and 3 together we have that $\Delta D_1 \geq 0$, which proves Inequality 1.

For the case of non-signatories (Inequality 2) we have that for phases 1 and 3, $\Delta D_2 = 0$, since here we are only concerned on the entering country. For these two phases, the entering country is the same for both *Cascadings*. Following the same reasoning of phase 2 in the previous paragraph, we get that $\Delta D_2 < 0$ for this phase. Putting the three phases together leads to $\Delta D_2 \leq 0$, which again proves Inequality 2 (note that Inequality 2 has the *Cascading* 1 and 2 inverted with respect to Inequality 1).

This last point proves that *Cascading* 1 implies *Cascading* 2. We iterate on the swapping process in r' until the recipient group becomes r^*, the least-green outsider, which finishes the proof.