Climate Coalitions and Punishments

Terry Eyland

Department of Economics, Bishop’s University, Sherbrooke, Canada
Email: teyland@ubishops.ca

Abstract

Studies that demonstrate that climate change is human induced are becoming more and more prevalent. Even though most world leaders are aware of this urgency and know that we must work at mitigating it quickly, little has been accomplished in terms of widespread participation in an International Environmental Agreement (IEA). The purpose of this paper is to create a link between studies on the use of border tax adjustments (BTAs) and coalition formation. The main contribution is that the punishment will be based on relative emissions between signatories and defectors. It is a structure that is more likely be accepted by the World Trade Organization (WTO) since it may be seen as fair due to the fact that if signatories and defectors emit the same amount of pollution then there will be no punishment. The main results indicate that this form of punishment may lead to small, partial, or full cooperation, depending on the parameter values. Additionally, at any equilibrium level, the signatories have a punishment structure that induces defectors to reduce their emissions by the same amount. In the end, this punishment may be seen as a credible threat because at equilibrium no punishment is imposed, yet if we remove the possibility of punishment it breaks down to a situation wherein no large coalitions are feasible.

Keywords

Environment, International Agreements, Non-Cooperative Game, Punishment Strategies

1. Introduction

Environmental issues are commonly of an international concern. The actions undertaken within a country are often felt by surrounding countries. Some issues such as climate change caused by the release of greenhouse gas (GHG) emissions in the atmosphere are international in nature because GHGs are transboundary pollutants. Hence, one unit of emissions released in one country versus another has the same impact on climate change. Viewing this obvious
link, countries have realized that this problem must be addressed through the use of an international environmental agreement (IEA) because there lacks a global government to determine, monitor and enforce reduction targets by every country.

When countries decide whether or not to join an IEA and set emission reduction targets, the policymakers in power first think of the potential gains and losses of doing so. To be implemented, it should be beneficial not only in the long-term (i.e. climate change mitigation), but also in the short term (i.e. protection of domestic jobs). Ideally to have a successful IEA, there would be incentives for the major polluting countries to participate. However, complete cooperation is typically unlikely because of the well-known free rider problem, whereby some individual countries are better off adopting lax environmental policies, while other countries strive to reduce emissions through stringent actions [1]. The reason for this problem is that reducing emissions entails some costs, borne only by that country, and in turn leads to a loss of competitiveness.

The use of restrictive trade policies has been suggested to level the competitiveness of the playing field by offering non-participating countries an incentive to join the coalition (see for example, [2], [3] and [4]). A common proposal is the implementation of a border tax adjustment (BTA) on the imports of energy-intensive goods from countries that do not commit to reducing their emissions as well as export rebates.¹ [2] argues that if a country is maximizing its welfare taking into account everyone’s emissions as environmental damages, it must either impose a carbon tariff or differentiate domestic carbon taxes to avoid large losses of competitiveness that bring about carbon leakage. An increasing number of jurisdictions around the world are seriously considering using carbon-motivated BTAs. The US, with the Waxman-Markey Bill, and the EU are currently contemplating the implementation of BTAs.²

The legality of a BTA has always been subject of discussion, yet many believe that it does not go against WTO rules (see for example [6], [7], [8], [9]). [6] proposed that BTAs should be imposed under the best available technology (BAT) rule to be in accordance with WTO regulations. Under this proposition, goods will be subject to the same “per unit” taxation level no matter the country in which they came from, thus simplifying greatly the informational requirements. [5] believes that the key to acceptance under the WTO is in how the BCA is designed and implemented.

Specific to climate change, many researchers have looked at finding ways to achieve the grand coalition to an international environmental agreement (IEA) (for a thorough review, see [10]). Overall, except in very specific cases (e.g.: dynamic games), researchers found the maximum coalition size to be of 2 or 3 [11]. [11] sees it as costlier to punish when one country pollutes more than the cooperative desired level. Costly punishments lead to non-credible or incredible threats wherein one country has no incentive to administer a punishment and

¹Also referred to as a border carbon adjustment in the literature dealing with the emissions trading scheme.
²For more on EU policies, see [5].
the other has no reason to believe that it will be punished. [12] argues that a punishment or negative incentive must be credible and severe for it to be effective. Ideally, it should be demonstrated that it is possible to punish a non-compliant country without much damage in the process.

Previous punishment structures relied on using more emissions as a form of punishment, which not only hurt the deviator but also the coalition making it a non-credible punishment. [13] looked at the evolution of signatories to an IEA on climate change. They developed a punishment structure imposed by signatories on to deviators which was linked to the number of signatories and the level of the pollution stock. Naturally as the pollution stock grew, the punishment would increase because it would become more pressing to limit emissions. Furthermore, the cost of punishment borne by the signatories would fall as more signatories enter. As in most papers looking at coalitions they use a stability concept introduced by [14], which states that for a coalition to be stable, no signatory can have an incentive to exit, nor does any non-signatory have an incentive to join the agreement. Their paper may be viewed as a good representation of reality, but for a punishment to be acceptable by the WTO it must be seen as fair. The question that seems unresolved is why would two deviators with different emission levels be subject to the same punishment? Or why can a signatory country who pollutes the same amount as a non-signatory inflict a punishment? The main shortfall in the noted article is that there is nothing that links punishment to relative emissions.

The purpose of this short essay is to expand on [13] by looking at the implications of introducing relative emissions, which gives it more of a BTA flavour. The main questions to be answered are

1. Will a punishment that depends on relative emissions cause “defectors” to emit less in order to suffer a smaller punishment?

2. If the answer to the first question is yes, does this mean that more countries will become “signatories” because they naturally emit a closer level to the signatories to avoid this emission dependent punishment?

The motivation here is to determine whether through the use of an emission indexed punishment there will be greater membership and lower global emissions. If yes, this would mean that we may no longer rely on an extensive IEA to succeed in fighting climate change, but rather a simple IEA with the permission of the WTO to impose a punishment and have everyone maximizing their self-interest.

The rest of the paper is structured as follows. The model will be described with key equations in Section 2. The game will be played in Section 3 and a sensitivity analysis will be tested in Section 4. Section 5 will conclude and provide some possible extensions.

2. The Model

As in [13], the current paper looks at a situation where we have \( N \) symmetrical countries that decide whether or not they want to become a “signatory” of an
IEA for the reduction of pollution. The set of signatory countries is denoted by $S$ and the members of $S$ maximize their aggregate welfare, where $s$ is the number of signatory countries, $|S| = s$, $s \geq 0$. The remaining $N - s$ countries choose not to join the agreement and are referred to as “defectors”, where $D$ denotes the set of defectors.

Production by all countries leads to benefits and pollution emissions. The benefits from production is represented by the following function

$$R(e_i) = e_i \left( b - \frac{1}{2} e_i \right),$$

where $e_i$ is the level of emissions by country $k \in S, D$ and $b$ is a positive scaling parameter.

For simplicity reasons and to better compare with [13], environmental damages are linear in the amount of emissions released by all the countries.

$$D(e) = d \sum_k (e_k),$$

where $d > 0$.

### 2.1. Punishment Function

In [13], all signatory countries agree to inflict a punishment to all non-signatory countries, and it is assumed to be proportional to the level of the pollution stock. It is difficult for such a punishment to be seen as “fair” if it is solely based on the current pollution stock and the number of signatories. In practice, it is challenging to justify a punishment from a signatory country if the non-signatory country has the same level emission reductions. Ideally, this punishment serves to penalize countries with laxer environmental policies as the Border Tax Adjustment (BTA) literature would suggest. Thus, if the emission reduction level is the same between signatory and non-signatory countries then there should be no punishment.

We introduce a component that makes the punishment smaller if the deviators work at reducing their emissions further.

$$\phi_i (s, e_i, e_j) = s \alpha (e_i - e_j),$$

where $e_i$ is the emission level of a defector country, $e_j$ is the emission level of signatory countries, and $\alpha \geq 0$ is a scaling parameter. The punishment function is such that, for a fixed level of emissions by the signatories, the punishment is increasing with the level of emissions by the defectors.\(^3\)

### 2.2. Self-Inflicted Cost Function

The self-inflicted cost will be modelled as a fraction of punishment as was done in [13]. The intuition here is that countries who choose to punish will incur a

\(^3\)Here, we assume that you can have a single country who is punishing all the others (a single “signatory”).

\(^4\)Note that if ever defectors emit less than signatories they can receive a side-payment instead of a punishment.
cost to do so. We assume that this net impact is a cost to the signatories, and that
this cost falls as more countries join the agreement.

\[ \omega_j(s,e_i,e_j) = \tau(N-s)\alpha(e_i-e_j) \]  

(4)

where \( \tau \) is the self-inflicted cost scaling parameter and \( \tau \geq 0 \). Participation in
an IEA has two extremes that must always be kept in mind. In the case where no
one wants to be part of the agreement and all act non-cooperatively, we need to
have no punishment nor any self-inflicted cost suffered. The other extreme case
is when we have full participation, where all maximize their welfare and take
into account the negative impact of their emissions on the environmental damages
suffered. In this case, we must also have no punishments nor self-inflicted cost.
The current model satisfies both of these extreme scenarios. It is the transition
between these two points that is interesting to examine.

As the number of signatories grows, the punishment which causes a reduction
in welfare to defectors, will increase. Since we are dealing with countries that are
standardized to being symmetric, comparing the amount of emissions of a typical
signatory to the defiator will allow us to discern how much more environmentally
friendly the signatories are relative to the defectors. Even though signatories
join an agreement, they should not be allowed to impose a punishment if their
emissions are equivalent to that of the defectors. The defiators emissions levels
are likely going to fall as they identify the link between their emissions and the
punishment. Their incentive to reduce emissions becomes greater as the number
of signatories, and thus the number of countries administering a punishment,
increases.

The key welfare equations of the model are

\[ W_j(\cdot) = e_j\left(b - \frac{e_j}{2}\right) - \tau\alpha\sum_{i \in D}(e_i-e_j) - d\left(\sum_{j \in S} e_j + \sum_{i \in D} e_i\right) \quad \forall j \in S, \]  

(5)

\[ W_i(\cdot) = e_i\left(b - \frac{e_i}{2}\right) - s\alpha(e_i-e_j) - d\left(\sum_{j \in S} e_j + \sum_{i \in D} e_i\right) \quad \forall i \in D. \]  

(6)

where \( \sum_{j \in S} e_j \) and \( \sum_{i \in D} e_i \) are respectively the total emissions of the signatories and
the defectors.

3. The Game

When the game is played for the signatory we are maximizing the welfare of a
signatory taking into account the negative externality it has on other signatory
countries, which is represented by the following,

\[ \max_{(e_j),j \in S} W^S_j = \sum_{j \in S} e_j\left(b - \frac{e_j}{2}\right) - \tau(N-s)\alpha(e_i-e_j) - d\left(\sum_{j \in S} e_j + \sum_{i \in D} e_i\right), \]  

(7)

Conversely the defector does not take into account any negative externality,

\[ \max_{(e_j),j \in D} W^D_j = e_i\left(b - \frac{e_i}{2}\right) - s\alpha(e_i-e_j) - d\left(\sum_{j \in S} e_j + \sum_{i \in D} e_i\right). \]  

(8)
Typically, we would assume that signatories would be the first movers and that defectors would then decide on their pollution levels. However, seeing that we have payoffs that are linear in the choice variable means that the reaction functions of signatories or defectors will not directly depend on the emission levels of the other group. Consequently, a sequential Stackelberg game will be the same as a simultaneous Nash equilibrium in this case. From taking the first-order conditions we get the following equations:

\[ e^S = b - ds + \tau (N - s)\alpha \]  
\[ e^D = b - d - s\alpha . \]  

Parameters such as a large \( b \) are chosen to ensure that \( e^S \geq 0 \) and \( e^D \geq 0 \). From here, we may be interested in the level of emissions of one group relative to that of the other. If we remove punishment and self-cost from the previous equations it is clear that as the number of signatories goes to a number greater than 1 the level of emissions from signatories will be smaller than the one from the defectors. Once we add punishment, we are inducing defectors to reduce emissions and incur a smaller punishment level. To answer the first question in the introduction, we may notice at this point that the emissions of defectors are based on the amount of signatories and their respective punishment. The more signatories join, the less defectors have an incentive to pollute. They may actually reduce their emissions by a larger percentage than signatories, thus giving them access to some form of compensation (i.e.: similar to the idea of the Clean Development Mechanism in the Kyoto Protocol).

\[ e^D - e^S = d (1 + s) - \alpha (s + N\tau - s\tau). \]

Based on this emission difference equation, it is interesting to note for what value of \( s \) the difference in emissions is equal to 0.5 One can remark how useful this piece of information is by looking at \( W^S \) and \( W^D \) above. Simply put, when emissions are the same for both groups we end up with no punishment or self-cost and the same firm profits and environmental damage, leading to the same welfare levels. Hence, for the parameter values used in the model, there should exist an \( s \) where \( e^D = e^S \).

\[ s^* = \frac{d + \alpha N\tau}{d + \alpha (\tau - 1)}. \]  

To ensure that this \( s^* \) is a positive number, we must first ensure that the parameter values satisfy \( \tau > \frac{d}{\alpha} \). If the punishment parameter \( \alpha \) is smaller than the damage parameter \( d \), this condition will be satisfied in all cases. However, if the punishment is greater than the environmental damage parameter, the self-inflicted cost of punishment must increase as the punishment increases or else there will not be a situation where the emissions are the same between both groups.

\(^5\) In a BTA model looking at strategic effects of BTA, [15] found a similar situation where the strategic effects would lead to the same emission level by both types of country.
If we believe that the self-inflicted cost will only be a fraction of the actual cost
\[ 0 \leq \tau \leq 1, \quad \alpha < \frac{d}{1 - \tau} \]
would be a sufficient condition to ensure that we have a positive \( s^* \). This condition simply requires the punishment parameter not to be too large relative to the environmental damage cost. It may still be larger because \( (1 - \tau) \leq 1 \) but should not be largely out of proportion.

Looking at Equations (9) and (10), we can see that in both cases as the number of signatories is higher the level of emissions will be lower. Before going too far on the discussion of what impacts \( s^* \), we must first look at the welfare for both groups. Inserting Equations (9) and (10) in Equations (5) and (6) we get the following:

\[
W^s = \frac{1}{2} \left( b^2 - 2bdN + d^2 \left( 2N + (2 + s)s \right) - 2\alpha d \left( N - s \right) \left( -s + (-1 + s)\tau \right) \right) + \alpha^2 \left( N - s \right) \tau \left( 2s + N\tau - s\tau \right) \]

\[
W^D = \frac{1}{2} \left( b^2 - 2bdN + d^2 \left( -1 + 2N + 2(-1 + s)s \right) + 2\alpha ds \left( N + s(-2 + \tau) - N\tau \right) + \alpha^2 s \left( s + 2N\tau - 2s\tau \right) \right) \]

and consequently,

\[
W^D - W^s = \frac{1}{2} \left( d \left( -1 + s \right) + \alpha s\tau - \alpha \left( s + N\tau \right) \right) \left( d \left( 1 + s \right) - \alpha \left( s - N\tau + s\tau \right) \right) \]

Setting \( W^D = W^s \) and solving for \( s^* \) we obtain two possible results:

\[
s^*_{W^D = W^s} = \frac{d + \alpha N\tau}{d + \alpha (\tau - 1)} \]

which is the same as obtained previously or

\[
s^*_{W^D = W^s} = \frac{d + \alpha N\tau}{\alpha - d + \alpha \tau} \]

To obtain a positive \( s^{**} \) we need \( \alpha > \frac{d}{1 + \tau} \) and \( \alpha \geq d \left( 1 + \frac{1}{N} \right) \) for \( s^{**} \leq N \). Thus, for a very large \( N \) we approximately need \( \alpha > d \). However, if we impose this condition after inserting this equilibrium \( s^{**} \) in the emissions difference equation,

\[
e^D - e^s \left( s^{**} \right) = -\frac{2(d - d\alpha N\tau)}{\alpha - d + \alpha \tau} \]

We obtain that at equilibrium the emissions of the defectors will be lesser than the signatories. Additionally, it requires that we have a punishment parameter greater than the actual environmental damage, which is unlikely going to be allowed by the WTO. We can thus eliminate this non-intuitive possibility.

After eliminating the possibility of \( s^{**} \), we may verify that \( s^* \) is indeed a stable equilibria. In coalition formation studies, a measure of stability uses \[14\] internal and external stability concept. Internal stability verifies that at this equilibrium there is no incentive for anyone that is part of the signatory to leave, and external stability ensures that no defector wants to join the signatory.
Looking locally first, we find that

\[ W^D - W^S (s^*) = 0 \]  \hspace{1cm} (17)

\[ W^D - W^S (s^* - 1) = \frac{1}{2} \left( -d^2 + \alpha + 2(\alpha - d)N\tau - \alpha\tau^2 \right) \]  \hspace{1cm} (18)

\[ W^D - W^S (s^* + 1) = \frac{1}{2} \left( 3d^2 + 2ad(-2+N\tau) - \alpha^2(-1+N\tau+\tau^2) \right). \]  \hspace{1cm} (19)

By the condition \( \alpha < \frac{d}{1-\tau} \) set earlier to have \( s^* > 0 \), we can confirm that this equilibrium is stable because when \( s < s^* \) we have \( W^S > W^D \) and for \( s > s^* \), we have \( W^D > W^S \).

Coming back to Equation (15), we had a condition to ensure that this number was positive and, if we have \( \alpha \leq \frac{d(N-1)}{N} \) we will have \( s^* \leq N \). We can notice that the value of \( \alpha \) that will cause \( s^* = N \) does not depend on \( \tau \) and this is simply because if \( s^* = N \), we have a situation where all countries are signatories and no self-cost is suffered. If \( \alpha > \frac{d(N-1)}{N} \) then the equilibrium coalition size is greater than the number of countries and there would be a convergence to the grand coalition.

4. Sensitivity Analysis

4.1. Comparative Static Analysis

Knowing that \( s^* \) is a stable equilibrium and that the greater the \( s \) the lower the total emissions, one may wonder what factors impact the value of \( s^* \). There are only four variables which may impact \( s^* \). Following is a comparative static analysis on how each influences \( s^* \).

\[ \frac{\partial s^*}{\partial N} = \frac{\alpha\tau}{d + \alpha(-1+\tau)} \]

\[ \frac{\partial s^*}{\partial d} = -\frac{\alpha(N\tau+1-\tau)}{(d + \alpha(\tau-1))^2} \]

\[ \frac{\partial s^*}{\partial \alpha} = \frac{d(1+(N-1)\tau)}{(d + \alpha(\tau-1))^2} \]

\[ \frac{\partial s^*}{\partial \tau} = \frac{\alpha(d(N-1)-\alpha N)}{(d + \alpha(\tau-1))^2}. \]

**Impact of \( N \) on \( s^* \):** If the number of symmetric countries in the model increases, the greater the equilibrium number of signatories based on the previous condition that \( d > \alpha \).

**Impact of \( d \) on \( s^* \):** As the environmental damage parameter becomes greater there are less signatories. This result can be explained by the fact that with greater environmental damage, the signatories will further reduce their emissions because they internalize this negative externality causing a greater burden on their firms'
profits, and hence reducing the number of countries having an incentive to become part of the signatory.

**Impact of \(\alpha\) on \(s'\):** If the punishment level rises, it would be less beneficial to be a defector and the number of *signatories* will increase. It was noted earlier that the punishment parameter should always be smaller than \(d\). This finding demonstrates that as the punishment rises, more countries will want to become signatories answering the second question from the introduction.

**Impact of \(\tau\) on \(s'\):** Fourth, the greater the self-cost parameter, the greater the number of *signatories* rendering the welfares equal, since *signatories* will now suffer a greater cost, it needs to be spread over a larger number of *signatories*.

These results have provided insights on how altering certain key variables allow us to get closer to the grand coalition. Ultimately, simply having a large enough punishment parameter (still smaller than the environmental damage parameter), may be a credible threat to obtain the grand coalition where no punishment is induced on anyone. Even though for certain parameter values we may not obtain the grand coalition, the punishment structure based on relative emissions substantially reduces the level of emissions of the *defectors* compared to the Business as Usual scenario with no punishment. The punishment now serves as a credible threat to inducing a reduction in emissions by all players. In the end, the goal may not necessarily to obtain the grand coalition but to induce a reduction by all.

### 4.2. Off Equilibrium Analysis

In equilibrium, we will always obtain the same amount of emissions for all countries, therefore no punishment or cost. An interesting observation is if we have an amount of *signatories* that is smaller than \(s'\), we have \(e^s > e^D\). This could be interpreted as the *signatories* emitting more to induce the *defectors* to join the coalition. However, this equilibrium is not stable because when that occurs we are now in a situation where \(W^S > W^O\) and more people want to become *signatories*. At an equilibrium coalition size, *signatories* do not emit more than *defectors*. Additionally, one may anticipate that some countries may not choose to adopt a Kyoto Protocol, but due to policies like the Kyoto Protocol Clean Development Mechanism, they may receive a subsidy if their emission reduction is greater than the *signatories*.

### 5. Conclusions

[13] had a punishment and self cost independent of the current choice of emissions. Introducing a punishment based on relative emissions changed the incentives of *defectors* and induced them to emit less than in the previous model. Based on what we know, this paper offers the first bridge between coalition formation and the BTA literature. Although it is a preliminary model, it offers interesting conclusions on the strategic impact of punishment based on relative emissions. First, a punishment that depends on relative emissions can cause the *defectors* to emit less in order to suffer a smaller punishment. Second, this new punishment
structure allows us to obtain a partial or the grand coalition. It also offers different solutions on how to increase the number of signatories, to potentially converge to the grand coalition where all pollute the cooperative ideal level of pollution.\(^6\)

The main deficiencies of this article can also be seen as future opportunities. For example, examining different functional forms such as a quadratic environmental damage function to have more interdependence in the reaction functions and to allow more of a sequential game. Moreover, an estimation of real world parameter values to determine the expected outcome would be very beneficial. Additionally, introducing dynamics or other markets may offer additional insights.

Acknowledgments

We thank the Editor and the referee for their comments. The research of Dr. Terry Eyland is funded by Bishop’s University. This support is greatly appreciated.

References


Economic Papers, 46, 878-894.


Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.
A wide selection of journals (inclusive of 9 subjects, more than 200 journals)
Providing 24-hour high-quality service
User-friendly online submission system
Fair and swift peer-review system
Efficient typesetting and proofreading procedure
Display of the result of downloads and visits, as well as the number of cited articles
Maximum dissemination of your research work

Submit your manuscript at: http://papersubmission.scirp.org/
Or contact tel@scirp.org