



Negotiating a uniform emissions tax in international environmental agreements

David M. McEvoy^{a,*}, Matthew McGinty^b

^a Department of Economics, Appalachian State University, Boone NC 28608, USA

^b Department of Economics, University of Wisconsin-Milwaukee, PO Box 413, Milwaukee, WI, 53201, USA



ARTICLE INFO

Article history:

Received 28 August 2017

Revised 31 May 2018

Accepted 1 June 2018

Available online 6 June 2018

JEL classification:

C7

D7

F5

H4

Keywords:

Global public goods

Uniform emissions tax

International environmental agreements

Externalities

ABSTRACT

A consensus appears to be emerging that a global carbon tax is the best policy for managing greenhouse gas emissions. Emissions tax systems are relatively straightforward, cost effective and can generate revenues used to offset other distortionary taxes. Moreover, recent theoretical research (Weitzman, 2014) has demonstrated that under some conditions the globally efficient tax rate can be implemented through a majority voting rule. We extend this area of research by examining a uniform emissions tax system in the framework of an international environmental agreement in which only countries that voluntarily participate are subject to the tax. We show that in the simplest situation in which countries have identical marginal benefit and cost functions, the largest stable agreement consists of two countries and the tax system has little impact on abatement levels. Our analysis highlights that by ignoring the participation decision and assuming commitment by all parties, the efficiency gains from a uniform emissions tax system are overstated.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Managing climate change is one of the most challenging public goods problems humanity has ever faced. While most of the scientific research suggests that collective welfare gains can be achieved by slowing climate change through greenhouse gas (GHG) emissions reductions, little progress has been made thus far. One of the reasons for such limited progress is that climate change is a social dilemma; one in which the collectively optimal scenario is at odds with what is individually optimal. The global community is better off restricting greenhouse gases while individual countries would prefer to free ride and let others abate for them. In the absence of an overarching authority that can impose regulations on countries with the goal of aligning collective and individual incentives (e.g., emissions taxes or permits), international environmental problems must be solved through international treaties. To date, international treaties on climate change have done little to solve the problem.

There is also continued debate regarding how to structure international policies to address climate change. Part of the debate is whether quantity based targets – like those fashioned under Kyoto – are the best approach. Many academics and policymakers have argued in favor of negotiating a uniform emissions price (i.e., a carbon tax) as a more desirable policy approach in

* Corresponding author.

E-mail addresses: mcevoydm@appstate.edu (D.M. McEvoy), mmcginty@uwm.edu (M. McGinty).

comparison to emissions permits or standards.¹ The idea of setting a single global emissions price to manage climate change was proposed early in the literature (Nordhaus, 1982; Hoel, 1991, 1992; Pearce, 1992; van der Ploeg and Zeeuw, 1992; Endres, 1997) and has gained traction more recently (Nordhaus, 2006, 2007; Cooper, 2010; Weitzman, 2014; Aldy, 2015; Cramton et al., 2015). With global externalities there are a number of arguments in favor of structuring negotiations on emissions taxes over quantities to regulate the environment (see Nordhaus (2006) for a detailed discussion). One argument centers around its simplicity; negotiating a single uniform emissions tax is more straightforward (with potentially fewer transactions costs) versus negotiating n different standards or permit allocations. There is also the potential for tax revenue (i.e., the “double dividend”) – that can be used to offset other distortionary taxes – that is not present with emissions standards or permits (if allocated freely). Less convincing, but still possible, is the argument that a uniform emissions tax has the appearance of being equitable since every country faces the same tax rate. This argument, of course, is only about the appearance of equity since the resulting level of emissions abatement will not be equitable. Given that marginal abatement costs are known, only the process and framing differ between abatement responsibilities under a price and quantity system.²

Perhaps the most compelling argument for regulating global emissions using prices over quantities is one put forth by Weitzman (2014).³ He points out that a uniform tax incorporates incentives for individuals to internalize the externality while these same incentives are not present in an individual quantity base system or a unilateral tax approach.⁴ This is because as each country (or player) considers its optimal tax rate they also take into account how the tax rate will affect the emissions choices of the others. Because each country faces the same tax, proposing a tax rate for itself is the same as proposing a tax rate for everyone. On the other hand, as a country proposes its individually optimal emissions cap, that cap only affects its own emissions choices and payoffs. Weitzman refers to this as a “countervailing force” embodied in a uniform tax system. He argues: “A binding uniform price of carbon emissions has a built in self-enforcing mechanism that countervails free riding.”

Weitzman goes on to model the endogenous determination of a global uniform emissions tax. In the model, the tax is negotiated by having each agent propose its individually optimal tax rate. The agents have heterogeneous marginal costs and benefits of abatement but will all ultimately face a uniform tax. The chosen tax will be determined through pairwise majority voting on all proposed taxes. The major result is that if the distribution of marginal benefits is symmetric (i.e., mean = median) then the chosen tax rate – by the median voter theorem – will equal the socially optimal tax rate. The implication is that under a majority rule voting system, even when countries vote selfishly for their optimal tax rate, the result is a socially optimal uniform emissions tax. This is a remarkable result. The desirable properties of such a system are further underscored when compared to a more complicated system in which n agents propose and negotiate n different emissions caps.

A uniform price on emissions, however, does not mitigate the fundamental free rider problem. At its core, climate change is a problem of collective action, one that requires individual countries to voluntarily opt into a treaty that regulates emissions. In the world modeled by Weitzman (2014), and considered informally by many since Hoel (1991) and Pearce (1992), countries are bound by the uniform emission tax. The model avoids the question of whether countries have an incentive to form such a treaty in the first place, and what level of participation a treaty could be expected to achieve. The fundamental problem with global externalities is that sovereignty precludes the possibility of a heavy-handed regulator imposing a tax on all countries. These are questions of *stability* in the context of an international environmental agreement (IEA). The literature on stable IEAs is well established, starting with an early paper by Barrett (1994) and extended by many (e.g., Ulph, 2004; Diamantoudi and Sartzetakis, 2006; Kolstad, 2007; McGinty, 2007; Finus and Maus, 2008; McEvoy and Stranlund, 2009; Eichner and Pethig, 2013). These papers develop game theoretic models of agreement formation and use the concepts of *internal* (no member wants to leave) and *external* stability (no non-member wants to join) to derive the equilibrium size of international environmental agreements.⁵ In these models members negotiate emissions abatement quantities, and in the basic model members agree to choose quantities to maximize their collective benefits. Even when considering a simplified world of symmetric countries, the models suggest international environmental agreements will consist of only a small subset of countries when there are substantial gains to cooperation. Nordhaus (2015) refers to this as the “small coalition paradox”.⁶ The intuition is that the stable equilibrium size is just large enough so that members and non-members are better off with an agreement. While perhaps counterintuitive, the research finds that coalitions are largest (smallest) when the benefits to cooperation are smallest (largest). This result, initially demonstrated by Barrett (1994), has been shown to be robust, persisting under various functional forms (e.g., Barrett, 2005), under heterogeneity with optimal transfers (e.g., McGinty, 2007), under increasing returns to shared environmental technology (e.g., Barrett, 2006), under dynamic extensions (e.g., Rubio and Ulph, 2007), and many others (e.g., Battaglini and Harstad, 2016, and see de Zeeuw (2015) for a recent survey of the IEA literature). Others have noted the shortcomings of top-down approaches to international environmental agreements (e.g., Edenhofer et al., 2015).

The objective of this paper is to consider the effectiveness of a uniform emissions tax mechanism when allowing countries to voluntarily participate in such a regime. The “countervailing force” of an emissions tax only works if other agents are subject

¹ The recent US carbon tax proposed by the Climate Leadership Council – positioned as a “conservative case for climate action” – highlights the broad interest in carbon pricing even across political parties (Feldstein et al., 2017).

² Carbon taxes have also been shown to outperform permits, both in theory and practice, under conditions of uncertainty (e.g., Hoel and Karp, 2002).

³ Weitzman expands on this model and compares it directly to a quantity based policy in Weitzman (2016).

⁴ For our purposes we use the terms *uniform* and *harmonized* tax interchangeably, meaning a single tax rate adopted by all countries.

⁵ Similar definitions of stability were used prior to Barrett (1994) by d'Aspremont et al. (1983) and Carraro and Siniscalco (1993).

⁶ Nordhaus (2015) explores a climate treaty structure that can potentially avoid the small coalition paradox by having the agreement members impose trade sanctions on non-members in order to increase participation.

to the tax. Specifically, we examine the stability of an international agreement that has its members negotiate a uniform price on emissions. Our approach links two strands of the literature: uniform emissions taxes and stable international environmental agreements. We adopt the uniform tax mechanism and functional forms from [Weitzman \(2014\)](#) and we model the participation decision and stability requirements from the IEA literature. Since Weitzman's principal finding is that the uniform emissions tax is efficient given that the marginal benefit intercepts are distributed symmetrically, we impose this (mean = median) by focusing on the simplest case in which countries are identical.

With identical countries, we show that stable coalitions can be no larger than two members under a uniform tax regime.⁷ We also demonstrate that IEA members (of any agreement size) will propose a uniform emissions tax that maximizes both individual and collective welfare of the members. Indeed, when countries are identical quantity-based IEAs and price-based IEAs lead to the same outcome. The important implication is that when countries are identical, only a small fraction of countries will join an IEA and reduce their emissions regardless of whether the policy regime is quantity based or price based. The advantages of the “countervailing force” of an emissions tax system unravel when participation in the tax regime is voluntary. The problem of internalizing externalities through voluntary arrangements has also been introduced outside of the IEA literature, and our findings relate to that research as well. [Dixit and Olson \(2000\)](#) add voluntary participation to a discrete-choice Coasian bargaining game and find that given nontrivial group sizes, free-riding incentives result in low participation rates and public goods are underprovided. As with our model, if voluntary participation is neglected and it is assumed that all players are willing to participate in the bargaining stage, then the players reach the efficient solution. Similarly, [Ellingsen and Paltseva \(2016\)](#) demonstrate that in groups with more than two players, incentives to free ride can prevent parties from voluntarily contracting to achieve efficiency gains through Coasian bargaining.

Our findings contribute to the ongoing discussion on how to design effective international environmental agreements. The promise of a uniform emissions tax as modeled in the literature has important implications for future treaties, and therefore the mechanism deserves close scrutiny. The remarkable result that a decentralized tax system can lead to efficient transboundary resource management is the product of a strong implicit assumption of reciprocity in a one-shot game; that is, each player is assumed to tax carbon at the globally determined price when others do so. This behavior, however, is neither a best-response nor a Nash equilibrium. We analyze the uniform emissions tax mechanism without making an assumption of full participation and obtain very different results.

The results we present are intuitive. Some may even criticize them as being predictable. But they are not obvious. The uniform emissions tax model differs from the standard IEA literature in a number of fundamental ways. Uniform taxes are considered instead of individual quantities. The tax is determined via a median voter rather than quantities chosen to maximize joint welfare. Of course, with identical countries all of the tax proposals are the same and equal to the median. The functional forms have non-zero marginal cost intercepts. Our model explicitly links both literatures to reveal that the encouraging properties of a uniform emissions tax largely hinge on an assumption of full participation.

Our approach begins by introducing benefit and cost functions that lead to the marginal functional forms proposed by [Weitzman \(2014\)](#). We then summarize the uniform emissions tax system proposed by Weitzman and reiterate his principal result. We then impose symmetry in the benefit and cost functions by specifying identical countries. To establish a baseline, we solve for Nash and optimal abatement levels. In the section that follows we develop the model of a uniform tax under an IEA with identical countries. We solve the game via backward induction, and begin by characterizing preferred tax rates and resulting abatement levels as a function of the number of IEA members. To help with the intuition in this stage we provide a numeric example. Then we solve for the equilibrium number of IEA members. We then offer some concluding remarks.

2. Model

We begin by considering a world with m countries, each indexed by $i = 1, 2, \dots, m$ that make decisions regarding emissions abatement levels. Country i 's abatement level is denoted as x_i and the aggregate abatement level is $X = \sum_{i=1}^m x_i$. We intentionally choose benefit and cost functions that lead to the marginal functional forms proposed by [Weitzman \(2014\)](#). It is important to note that the m players in our model are specified as countries while the m players in [Weitzman \(2014\)](#) are individuals. While there are no practical differences in the way the two models are analyzed, there are differences in the interpretations. We assume a country acts as a single player to maximize its individual payoff. Presumably the country is acting on behalf of its citizens and perfectly embodies their preferences.

The benefit to country i from aggregate abatement X is

$$B_i(X) = b_i X - \frac{\beta}{2} X^2, \quad (1)$$

and the marginal benefit of abatement, from [Weitzman \(2014\)](#), is

$$B'_i(X) = b_i - \beta X. \quad (2)$$

⁷ This result is consistent with the simultaneous-choice IEA model with members choosing quantities to maximize their collective welfare (see [Finus and Rundshagen, 2003](#)).

The cost of abatement depends only on individual abatement x_i and is

$$C_i(x_i) = c_i x_i + \frac{\gamma}{2} x_i^2, \quad (3)$$

thus the marginal abatement cost, from Weitzman (2014), is

$$C'_i(x_i) = c_i + \gamma x_i. \quad (4)$$

Note that the linear marginal benefit and cost functions have identical slopes for each of the m countries. The heterogeneity in the marginal functions is captured entirely through variation in the intercepts. Therefore, these linear formulations imply a ranking of marginal benefits and costs by b_i and c_i , respectively, when abatement is zero.

2.1. Weitzman's global uniform emissions tax

In this section we introduce the uniform emissions tax (or emissions price) system proposed by Weitzman (2014), and briefly highlight his main results. Reiterating these findings is important as they will serve as our main points of comparison when introducing the tax system within an international environmental agreement. Under the uniform tax, all m parties are subject to an identical per-unit emissions price that is decided by pair-wise majority voting.⁸

Let p denote the uniform international emissions tax that each country faces as a result of a pair-wise majority vote on the set of individually proposed taxes, where p_i indicates the preferred tax of country i . Given a uniform tax, the response of each country is to choose abatement by setting the tax equal to its marginal abatement cost from (4). Therefore we have $p = c_i + \gamma x_i$ and when rearranged yields country i 's abatement as a function of the tax

$$x_i = \frac{p - c_i}{\gamma}. \quad (5)$$

Since $X = \sum_{i=1}^m x_i$ we can form the expression for aggregate abatement as a function of the uniform tax

$$X = \frac{mp - \sum_{i=1}^m c_i}{\gamma}. \quad (6)$$

When proposing its individually optimal tax rate, country i considers a world in which $p_i = p$. Using this relationship, we can substitute the individual and aggregate levels of abatement as a function of p_i from Eqs. (5) and (6) into the individual marginal benefit and cost functions in Eqs. (2) and (4), respectively. Setting $B'_i(X(p_i)) = C'_i(x_i(p_i))$ and solving we get

$$p_i = \frac{m\gamma b_i + m\beta \sum_{i=1}^m c_i}{\gamma + m^2\beta},$$

which is country's i 's individual optimal emissions tax if imposed on all m countries. Note that Weitzman expresses this individual optimum as

$$p_i = kb_i + k', \quad (7)$$

where $k \equiv \frac{m\gamma}{\gamma + m^2\beta}$ and $k' \equiv \frac{m\beta \sum_{i=1}^m c_i}{\gamma + m^2\beta}$. We can now compare the individual optimal with the social optimal emissions tax. The aggregate payoff function for all m countries is $\Pi = \sum_{i=1}^m B_i(X) - \sum_{i=1}^m C_i(x_i)$. If we substitute the individual and aggregate levels of abatement as a function of price from Eqs. (5) and (6) into the aggregate payoff function for all m countries and solve for p we get

$$p^* = k\bar{b} + k', \quad (8)$$

where the mean marginal benefit intercept is $\bar{b} = \frac{\sum_{i=1}^m b_i}{m}$. From (7) it is clear that the preferred ordering of carbon taxes depends on the intercepts of the marginal benefits. If we consider pair-wise majority voting on the preferred value of p , by the median voter theorem the outcome will be the median value of p , which is denoted by \hat{p} . We can express this as

$$\hat{p} = k\hat{b} + k', \quad (9)$$

where \hat{b} is the median value of the marginal benefit intercept b_i . A comparison of (8) with (9) indicates that the result of the majority vote on a uniform emissions tax will be equal the social optimum if $\bar{b} = \hat{b}$. In other words, if the distribution of marginal benefits (the intercepts) is symmetric then the uniform emissions tax chosen by majority vote will be optimal. However, as Weitzman (2014) demonstrated, the further the distance between the average and median values of the marginal benefit intercepts (i.e., how skewed the distribution is), the farther the median preferred tax is from the socially optimal tax.

⁸ Weitzman (2014) appeals to the median voter theorem in order to specify the winning tax in a pair-wise majority voting system as the median proposal. Each agent has single-peaked preferences (i.e., a single ideal tax rate) and is voting on only one dimension, and therefore the outcome is independent of the order of voting.

3. Identical countries and abatement levels

Our primary goal is to consider the uniform emissions tax described in the previous section in the context of an international environmental agreement. The principal finding of Weitzman (2014) is that the uniform tax is Pareto efficient if the distribution of marginal benefit intercepts is symmetric (median = mean). Given this result, we proceed with our analysis using simplified functional forms in which the marginal benefit and cost intercepts are identical for all m countries. That is, we consider a world of identical countries by setting $b_i = b$ and $c_i = c$. In this case $\bar{b} = \hat{b}$, and therefore $\hat{p} = p^*$.

Before introducing the international environmental agreement with the uniform tax, we begin by establishing a baseline level of emissions abatement that would occur under unilateral management (i.e., the Nash equilibrium abatement levels). Using Eqs. (1) and (3) in our world of homogeneous countries, i 's payoff as a function of abatement is $\pi_i = B_i(X) - C_i(x_i)$ or

$$\pi_i = bX - \frac{\beta}{2}X^2 - cx_i - \frac{\gamma}{2}x_i^2. \quad (10)$$

Each country chooses its abatement x_i to maximize (10) which leads to the following first-order condition⁹:

$$\frac{\partial \pi_i}{\partial x_i} = b - \beta X - c - \gamma x_i = 0. \quad (11)$$

To simplify terms, let's define the net marginal benefit at zero abatement as $\phi \equiv b - c$, which is assumed positive otherwise the problem is trivial and abatement is zero. Solving (11) for x_i results in country i 's best-response function

$$x_i^{br} = \frac{\phi - \beta X_{-i}}{\gamma + \beta}, \quad (12)$$

where $X_{-i} \equiv \sum_{j \neq i} x_j$. The best-response slopes for all m countries are linear and show that abatement levels are strategic substitutes; that is,

$$\frac{\partial x_i^{br}}{\partial X_{-i}} = \frac{-\beta}{\gamma + \beta} \in (0, -1) \quad \forall i \in M. \quad (13)$$

Note from (13) that the slope of the best response function can be interpreted as the degree of carbon leakage, and is identical for all m countries. It is how much others reduce abatement in response to a unit increase in abatement by i . As γ approaches zero the best-response slopes approach -1 and there is complete carbon leakage. Likewise, as γ increases the slope of the best-response curve approaches zero and the amount of carbon leakage decreases. The slope of the marginal benefit function β has the opposite effect on leakage. As β decreases carbon leakage decreases and as β increases carbon leakage increases.

The simultaneous solution to the best-response functions results in individual Nash-equilibrium abatement levels

$$x^{ne} = \frac{\phi}{\gamma + \beta m}. \quad (14)$$

Aggregate abatement at the Nash equilibrium is

$$X^{ne} = \sum_{i=1}^m x_i^{ne} = \frac{m\phi}{\gamma + \beta m}. \quad (15)$$

Next we derive optimal abatement levels for the set of m countries. The goal here is to demonstrate that Nash abatement levels are suboptimal and there is a collective incentive to increase abatement. Together, the Nash and optimal abatement levels provide the bookends for agreements with partial participation. The global payoff from emissions abatement is

$$\Pi = \sum_{i=1}^m bX - \frac{\beta}{2} \sum_{i=1}^m X^2 - \sum_{i=1}^m cx_i - \frac{\gamma}{2} \sum_{i=1}^m x_i^2. \quad (16)$$

The social optimum is obtained by maximizing Π by choosing abatement x_i . The first-order condition is

$$\frac{\partial \Pi}{\partial x_i} = mb - m\beta X - c - \gamma x_i = 0, \quad (17)$$

which yields

$$x^* = \frac{mb - c}{\gamma + \beta m^2}. \quad (18)$$

The expression x_i^* is the optimum abatement level of country i , and aggregate optimal abatement is

$$X^* = \sum_{i=1}^m x_i^* = \frac{m(mb - c)}{\gamma + \beta m^2}. \quad (19)$$

⁹ The second-order condition is $-\beta - \gamma < 0$.

A comparison of the aggregate Nash equilibrium level of abatement from Eq. (15) with the aggregate optimal level of abatement from Eq. (19) reveals that $X^* - X^{ne} > 0$ if $(m-1)m[\gamma b + \beta mc] > 0$, which is true for all $m > 1$ and therefore we have a familiar social dilemma in which individually rational choices are not collectively rational.

4. IEAs with uniform emissions taxes

In Section 2.1 we considered a uniform emissions tax that is imposed on all m countries as the result of pair-wise majority voting. Weitzman (2014) demonstrates that when marginal benefits are distributed symmetrically, the result of the vote will be an optimal global emissions tax. In this section we examine the uniform emissions tax policy in the context of an international environmental agreement. The critical feature is that only countries that voluntarily become members to the IEA are part of the uniform tax system. As sovereignty precludes the possibility of requiring all m nations to participate in a uniform tax system, it is important to evaluate the mechanism in the context of an IEA.

We consider a world in which the m countries decide independently and simultaneously whether they want to join an IEA. Those that join - the signatories - opt into the uniform emissions tax system described in the previous section, while any non-signatories make their abatement decisions unilaterally. Specifically, the signatories individually propose their preferred emissions taxes and the uniform tax is chosen by pair-wise majority vote.

Before moving forward we want to point out how the institution we consider fits in with the existing literature. The uniform tax policy is the one Weitzman (2014) considers formally, but has been proposed by many others (e.g., Nordhaus, 2006; Cooper, 2010). It is adapted here by allowing only willing participants to be governed by a uniform tax under a voluntary agreement. We follow the IEA literature in that only signatories are subject to an emissions abatement policy and we determine the equilibrium number of signatories by relying on the internal and external stability conditions common in the literature. However, while signatories in our game choose emission taxes, the existing IEA literature has signatories choosing abatement quantities to maximize their collective welfare (e.g., Barrett, 1994). This departure from the standard setup is of course necessary, but it is also important to point out that in our simple case in which all countries are identical, the two approaches (prices versus quantities) lead to the same outcome. Finally, we also depart from the very common assumption in the IEA literature that abatement decisions are made sequentially with signatories acting as Stackelberg leaders (as in Barrett (1994)). To stay true to the uniform emission tax mechanism envisioned by Weitzman (2014) both signatories and non-signatories make their decisions simultaneously in a game with Cournot timing.¹⁰

Our game has two stages. In the first stage, countries decide individually whether to join the agreement as signatories. In the second stage, signatories participate in the uniform emissions tax system described in the previous section while non-signatories make their abatement decisions unilaterally. We begin by analyzing the second stage in order to derive emissions taxes and abatement levels for both signatories and non-signatories as a function of the number of members.

4.1. Signatory and non-signatory best responses

Let s denote the number of signatories to the IEA and each proposes an individually optimal emissions tax p_i^s that, if adopted, is uniformly imposed on all s signatories. Non-signatories are not subject to the uniform tax and choose an abatement level to maximize individual payoffs which can also be expressed as a function of their individually preferred tax, p_j^t .

Each country's payoff is determined by Eq. (10), and from (5) each signatory abates $x_i^s = \frac{p_i^s - c}{\gamma}$ and each nonsignatory abates $x_j^t = \frac{p_j^t - c}{\gamma}$, and therefore the payoff function for a signatory in terms of emissions taxes is

$$\pi_i^s = b \left(\frac{s(p_i^s - c)}{\gamma} + \sum_{j=s+1}^m \frac{p_j^t - c}{\gamma} \right) - \frac{\beta}{2} \left(\frac{s(p_i^s - c)}{\gamma} + \sum_{j=s+1}^m \frac{p_j^t - c}{\gamma} \right)^2 - c \left(\frac{p_i^s - c}{\gamma} \right) - \frac{\gamma}{2} \left(\frac{p_i^s - c}{\gamma} \right)^2.$$

Each signatory maximizes π_i^s by choosing p_i^s , and setting the first-order condition to zero yields

$$\frac{sb}{\gamma} - \frac{s\beta}{\gamma} \left(\frac{s(p_i^s - c)}{\gamma} + \sum_{j=s+1}^m \frac{p_j^t - c}{\gamma} \right) - \frac{c}{\gamma} - \left(\frac{p_i^s - c}{\gamma} \right) = 0.$$

Given symmetric first-order conditions for signatories we know $p_i^s = p_s$, and recognizing the symmetry of non-signatories $\sum_{j=s+1}^m = (m-s)$ and $p_j^t = p_t$, we can solve for p_s to derive a signatory's best response function

$$p_s^{br} = \frac{s[b\gamma - \beta[(m-s)p_t - mc]]}{s^2\beta + \gamma}. \quad (20)$$

¹⁰ Finus and Rundshagen (2003) analyzed a quantity-based IEA given Cournot timing. Later in this section we confirm that our stability results are consistent with theirs given homogeneous agents.

Note that if all countries are signatories then $s = m$ and this becomes

$$p_{s|m}^{br} = \frac{m[b\gamma + \beta mc]}{m^2\beta + \gamma} = p^*,$$

which is equal to the socially optimal carbon tax from (8) (and from Weitzman's Eq. (15)) when countries have identical marginal benefits and costs.

Meanwhile, non-signatories choose their carbon tax p_j^t to maximize their individual payoff

$$\pi_j^t = b \left(\sum_{i=1}^s \frac{p_i^s - c}{\gamma} + \sum_{j=s+1}^m \frac{p_j^t - c}{\gamma} \right) - \frac{\beta}{2} \left(\sum_{i=1}^s \frac{p_i^s - c}{\gamma} + \sum_{j=s+1}^m \frac{p_j^t - c}{\gamma} \right)^2 - c \left(\frac{p_j^t - c}{\gamma} \right) - \frac{\gamma}{2} \left(\frac{p_j^t - c}{\gamma} \right)^2,$$

and setting the first-order condition to zero results in

$$\frac{b}{\gamma} - \frac{\beta}{\gamma} \left(\sum_{i=1}^s \frac{p_i^s - c}{\gamma} + \frac{p_j^t - c}{\gamma} + \sum_{k \neq j, s+1}^m \frac{p_k^t - c}{\gamma} \right) - \frac{c}{\gamma} - \frac{2\gamma}{2\gamma} \left(\frac{p_j^t - c}{\gamma} \right) = 0.$$

As above, recognizing the symmetry and simplifying terms results in the non-signatory best-response function

$$p_t^{br} = \frac{b\gamma - \beta[sp_s - cm]}{\beta(m-s) + \gamma}. \quad (21)$$

Note that if $s = 0$ then the non-signatory best-response function reduces to the non-cooperative Nash equilibrium

$$p_{s|s=0}^{br} = \frac{b\gamma + \beta cm}{\beta m + \gamma} = p^{ne},$$

where abatement is $x_i^{ne} = \frac{p^{ne} - c}{\gamma}$.

4.2. Equilibrium taxes and abatement

The equilibrium emissions taxes for signatories and non-signatories are determined by the simultaneous solution to the best response functions (20) and (21). The result leads to our first proposition.

Proposition 1. Consider an IEA with a uniform emissions tax. The Nash equilibrium tax rates for (i) signatories and (ii) non-signatories are:

$$p_s = \frac{s[b\gamma + \beta mc]}{\beta(s^2 + m - s) + \gamma}$$

$$p_t = \frac{b\gamma + \beta mc}{\beta(s^2 + m - s) + \gamma}.$$

Proof. Derivation in Appendix A.1. ■

Note that the equilibrium tax for non-signatories is strictly decreasing in the number of signatories to the IEA, while the tax for signatories is non-monotonic in IEA membership. To help illustrate this relationship, Table 1 in the following section links taxes and the number of signatories for a given set of parameters. Equipped with our expressions for equilibrium emissions taxes for signatories and non-signatories, it is straightforward to solve for abatement levels and payoffs. Substituting the equilibrium tax of p_s from Proposition 1 into the abatement expression from (5) we can solve for abatement (derivation in Appendix A.2)

$$x_s = \frac{\gamma(sb - c) + c\beta(m - s)(s - 1)}{\beta\gamma(s^2 + m - s) + \gamma^2}. \quad (22)$$

Note that if $s = m$ then signatory abatement is equal to the social optimum

$$x_{s|m} = \frac{\gamma mb - c}{\beta m^2 + \gamma} = x^*.$$

Similarly, substituting p_t from Proposition 1 into abatement from (5) results in non-signatory abatement as a function of the number of signatories (derivation in Appendix A.2)

$$x_t = \frac{\gamma(b - c) - c\beta s(s - 1)}{\beta\gamma(s^2 + m - s) + \gamma^2}. \quad (23)$$

Note that if $s = 0$ then (23) reduces to non-cooperative Nash abatement

$$x_{t|s=0} = \frac{b - c}{\beta m + \gamma} = x^{ne}.$$

Table 1
Emissions taxes, abatement levels and payoffs for signatories and nonsignatories.

s	p_s	p_t	x_s	x_t	X	π_s	π_t	Π
0	–	50.00	–	5.00	50.00	–	3600.00	36,000.00
1	50.00	50.00	5.00	5.00	50.00	3600.00	3600.00	36,000.00
2	90.00	45.00	10.00	4.38	55.00	3487.50	3867.19	37,912.50
3	112.50	37.50	12.81	3.44	62.50	3512.11	4215.23	40,042.97
4	120.00	30.00	13.75	2.50	70.00	3656.25	4500.00	41,625.00
5	118.42	23.68	13.55	1.71	76.32	3849.31	4690.72	42,700.14
6	112.50	18.75	12.81	1.09	81.25	4039.45	4808.50	43,470.70
7	105.00	15.00	11.88	0.63	85.00	4204.69	4879.69	44,071.88
8	97.30	12.16	10.91	0.27	87.84	4340.62	4923.05	44,571.04
9	90.00	10.00	10.00	0.00	90.00	4450.00	4950.00	45,000.00
10	83.33	–	9.17	–	91.67	4537.50	–	45,375.00

Given any level of IEA membership, aggregate abatement is $X = sx_s + (m - s)x_t$, which, when using (22) and (23), can be simplified to

$$X = \frac{sb(s-1) + m(b-c)}{\beta(s^2 + m - s) + \gamma}. \quad (24)$$

If all countries join the IEA then $s = m$, $X = mx_s = X^*$ and the uniform emissions tax is socially optimal, while if $s = 0$ or $s = 1$ then $X = mx^{ne} = X^{ne}$. Substituting the abatement levels from the two extreme cases where $s = m$ and $s = 0$ into the payoff function, we can solve for the Nash equilibrium and optimal payoffs.

$$\pi^{ne} = \frac{(b-c)[(\beta m^2 - \gamma)(b+c) + 2m(b\gamma - \beta c)]}{2(\beta m + \gamma)^2}$$

$$\pi^* = \frac{(bm - c)^2}{2(\beta m^2 + \gamma)}.$$

Next we demonstrate how aggregate abatement depends on the number of signatories.

Proposition 2. Under a uniform emissions tax, global abatement is strictly increasing in the number of signatories to a non-trivial (i.e., $s > 1$) IEA.

Proof.

$$\frac{\partial X}{\partial s} = \frac{b[2s-1][\beta(s^2 + m - s) + \gamma] - [sb(s-1) + m(b-c)][\beta(2s-1)]}{[\beta(s^2 + m - s) + \gamma]^2}$$

which is positive if

$$[2s-1][b\gamma + cm\beta] > 0.$$

In the case where $s = 1$ then we are back to unilateral management and the previously established non-cooperative Nash equilibrium. For all non-trivial coalitions in which $s > 1$, $\frac{\partial X}{\partial s} > 0$.¹¹ ■

4.3. A numeric example

In Table 1 we use a set of parameters to illustrate the relationship between the number of signatories, emissions taxes, abatement levels and payoffs for both signatories and non-signatories. In the table we fix $m = 10$, $b = 100$, $c = 10$, $\beta = 1$ and $\gamma = 8$. At $s = 0$ or $s = 1$ the IEA is of trivial size and reverts to the non-cooperative Nash equilibrium emissions price ($p^{ne} = 50$) and abatement ($x^{ne} = 5$). When $s \geq 2$, signatories choose tax rates that go beyond the non-cooperative Nash levels, and their abatement levels increase as a result. Note that at small IEA sizes of $s = 2$ and $s = 3$, signatories earn lower payoffs (3,487.50 and 3,512.11, respectively) compared to the payoffs they earn under the non-cooperative Nash equilibrium ($\pi^{ne} = 3,600$). This occurs because abatement levels are strategic substitutes and non-signatories reduce abatement relative to the Nash equilibrium. However, as s increases so do the payoffs to signatories and eventually (i.e., $s \geq 4$) signatories earn higher payoffs as members to the IEA compared to the non-cooperative baseline. Note, however, that non-signatories always earn strictly higher payoffs than signatories (when two or more exist). When $s = 10$, the emissions taxes and abatement levels are at the social optimum ($p^* = 83.33$ and $x^* = 9.17$) and aggregate payoffs are maximized ($\Pi^* = 45,375$).

¹¹ This finding may not hold in related contexts. For example, Goeschl and Perino (2017) present a model of an IEA where signatories face a uniform price for a new abatement technology and demonstrate that total abatement can be decreasing in the number of signatories.

4.4. Coalition stability

To solve for the equilibrium number of signatories we rely on the *internal* and *external* stability conditions common in the IEA literature (e.g., Barrett, 1994). An IEA is considered internally stable if no signatory finds it individually beneficial to leave the agreement. An IEA is externally stable if no non-signatory finds it beneficial to join the agreement. A signatory in IEA size s compares its payoff as a signatory with that of a non-signatory with $s - 1$ remaining members. An IEA is internally stable if

$$\pi_i^s(s) - \pi_i^t(s - 1) \geq 0, \quad (25)$$

and externally stable if

$$\pi_i^s(s + 1) - \pi_i^t(s) < 0. \quad (26)$$

An IEA is considered stable if it is both internally and externally stable.

Since a stable IEA needs to be both internally and externally stable, we focus on the largest internally stable IEA. With homogeneous countries, if an enlarged IEA with $s + 1$ is not internally stable then no non-member wants to join and an IEA of size s must be externally stable. The example in Table 1 illustrates that external stability holds for all $s \geq 1$ since $\pi_i^s(s) > \pi_i^s(s + 1)$. We now turn to analyzing the internal stability requirement. Aggregate abatement with s signatories is $X(s)$, which becomes $X(s - 1)$ if i becomes a non-signatory. Abatement as a signatory with s IEA members is $x_i^s(s)$ and abatement if i were to become a non-signatory is $x_i^t(s - 1)$. The internal stability (IS) condition in terms of abatement is

$$bX(s) - \frac{\beta}{2}[X(s)]^2 - cx_i^s(s) - \frac{\gamma}{2}[x_i^s(s)]^2 - \left[bX(s - 1) - \frac{\beta}{2}[X(s - 1)]^2 - cx_i^t(s - 1) - \frac{\gamma}{2}[x_i^t(s - 1)]^2 \right] \geq 0. \quad (27)$$

When substituting in the expressions for abatement, the equation above results in a sixth order polynomial with six variables. Previous research has resorted to simulations in order to characterize stable coalition sizes because finding analytical solutions to similar stability conditions was intractable (e.g., Barrett, 1994; Gelves and McGinty, 2016). However, we have identified an approach that allows us to demonstrate analytically that the largest stable coalition size is $s = 2$. We start by setting the IS condition equal to zero and solving for the critical value of m as an implicit function of the number of signatories and the parameters. If this critical value is such that, for integer values, $2 \leq s \leq m$, then the IS condition can be satisfied for a given IEA size. The critical value does not depend on the intercepts b and c when countries are identical. Denote this critical value as m^* , and is the positive root

$$m^*(s, \beta, \gamma) = \frac{-\beta s(s^2 - 2s - 1) - \gamma(s - 1) + 2\sqrt{(\beta + \gamma)(\beta s^2 + \gamma)}}{\beta(s + 1)}. \quad (28)$$

Note that the IS condition is zero if $m = m^*$. Moreover, the range of $m^* < s$ is not relevant because the number of signatories cannot exceed the critical number of countries. We focus our attention on the relevant range in which $m^* \geq s$.

Proposition 3. *Under a uniform emissions tax a stable IEA never has more than two members.*

Proof. The proof proceeds as follows (i) we show that an IEA of size $s = 2$ can be stable under certain conditions, and (ii) an IEA of size $s \geq 3$ cannot be stable under any condition. To accomplish (ii), we show that the IS condition for an IEA of size $s \geq 3$ cannot change sign for any positive parameters. This means that if an IEA is not internally stable for any $s \geq 3$ then it is unstable for all $s \geq 3$. (iii) We then show that the grand coalition $s = m \geq 3$ is unstable, hence all IEAs with more than two members are unstable. Details are in Appendix A.3. ■

To summarize, there is never a stable IEA with more than two members when countries have identical marginal benefit and cost functions.¹² This result has important implications. The potential efficiencies associated with a global uniform emissions tax begin to unravel when the mechanism is positioned as part of an international environmental agreement. In our analysis we explored the simplest situation that satisfies Weitzman's efficiency criteria - one in which m countries have identical benefit and cost functions - and we find that the only possibility for a stable IEA is two members.

Moreover, we show that a stable IEA of size $s = 2$ is possible but not guaranteed. Consider again the numeric example from Table 1. With $\beta = 1$ and $\gamma = 8$ the critical value m^* is equal to 4.93 when $s = 2$. Given these parameters, an agreement of size $s = 2$ is stable only in the range of $2 \leq m \leq 4$. The implication is that m needs to be sufficiently small for a stable two member IEA to exist for this parameter pair.

Proposition 4. *Quantity-based self-enforcing agreements and uniform emissions tax agreements will result in the same stable agreement size and equilibrium abatement levels.*

¹² Finus and Rundshagen (2003) using different functional forms also show that the largest stable IEA is of two members given Cournot timing of abatement decisions by signatories and non-signatories. In their model, however, signatories choose abatement levels to maximize joint payoffs.

Proof. The proof proceeds as follows. Following the IEA literature we consider an agreement in which signatories maximize aggregate payoffs by choosing abatement quantities. Non-signatories maximize their individual payoffs by choosing abatement. We show that the Nash equilibrium is identical to the results from a uniform emissions tax in (22) and (23). See Appendix A.4. ■

In summary, we find that an IEA in which its members adopt a uniform emissions tax is not expected to motivate meaningful participation, and abatement levels are not expected to increase much beyond the non-cooperative baseline. Although specifying identical functional forms for the m countries is overly simplistic, the assumption ensures that the symmetry requirement of Weitzman's result is satisfied. That is, if an IEA of size $s = m$ were to form then the members would self-impose the socially optimal uniform emissions tax. However, when many players are involved the grand coalition is not stable and the free-rider problem is not overcome. We have demonstrated that with symmetric countries an IEA is not expected to improve global efficiency measures for emissions abatement.

5. Conclusion

The prospect of managing climate change with an internationally harmonized emissions tax was established early in the economics literature. Lately, the arguments in favor of a price based system over a quantity based system appear to be gaining more traction. While there are many reasons why prices may be preferable to quantities, perhaps the most compelling is the one entertained by Cramton et al. (2015) and formalized by Weitzman (2014, 2016). Weitzman (2014) considers a system in which agents propose their individually optimal taxes and finds that under some conditions the socially optimal tax can be the outcome of majority voting. Specifically, provided that the distribution of marginal abatement benefits is symmetric, the median preferred tax will equal the optimal tax. The implication of the result is that even when countries act selfishly in their proposal for a uniform emissions tax, the result may be efficient. Weitzman refers to this as a “countervailing force” of a democratically-determined price based regulation.

International environmental problems, however, are problems of collective action. The aforementioned tax system circumvents the underlying social dilemma posed by climate change by requiring all affected parties to participate in the emissions tax system. Of course sovereignty prevents the possibility of forced participation. Our objective with this paper is to analyze the effectiveness of the uniform price system within the context of an international environmental agreement. We connect two strands of literature: one focused on uniform emissions taxes and the other on self-enforcing IEAs. We combine the tax mechanism and functional forms from Weitzman (2014) with the participation decision and stability requirements from the IEA literature.

We show that in the simplest scenario in which countries have identical marginal benefits of abatement (and therefore satisfying the efficiency criteria in Weitzman (2014)) agreements with more than two countries cannot be stable. With a global environmental problem like climate change this result suggests that an IEA based on a uniform emissions price is unlikely to improve efficiency compared to unilateral management. The countervailing force of an emissions tax is absent when only a small subset of countries participate. Although we do not formally consider heterogeneous countries in our IEA model, some of the implications are immediately clear. Stability aside, if the distribution of marginal benefits of IEA members is asymmetric, then the members should not expect to fully internalize the externality. That is, for any size agreement of asymmetric countries the outcome of the vote is a tax that is not collectively rational for the members.

Our paper contributes to the ongoing research on comparing prices versus quantities in mitigating global environmental problems. However, it is important to note that the prices versus quantities comparison made in this paper is notably different from the comparison made in much of the literature beginning with Weitzman (1974). This is because we are not comparing taxes with permits. Rather, we are comparing the results from the well-established IEA literature in which members choose quantities to maximize their collective payoffs with a novel IEA in which members propose their individually optimal tax rates. With identical countries, the two approaches are effectively the same. One possible criticism of this research is that we only consider countries with identical payoffs. Although exploring heterogeneity is certainly an avenue for future research, our objective is to take the institution considered recently in the literature and evaluate it when participation is voluntary. We chose the most straightforward environment in which countries have identical payoff functions in order to satisfy the aforementioned efficiency requirements of a negotiated uniform emissions tax.

Like much of the IEA literature our results are largely pessimistic. If stable IEAs exist with a uniform tax they will be very small and largely ineffective. Of course our modeling approach ignores a multitude of other incentives countries may have to join an international agreement. Research shows that a willingness to cooperate is not just based on evaluating own payoffs, but can be influenced by preferences toward equity and responsibility. There may also be positive reputation effects from cooperating on one global initiative that can spill over to other policy domains. In some cases countries cooperate in order to demonstrate a leadership role in the international community. Acknowledging the limitations of our model, however, does not dilute the fundamental result of the paper; that is, by ignoring the participation decision and assuming commitment by all parties, the efficiency gains from a uniform emissions tax system are likely overstated.

Acknowledgements

The authors would like to thank John Stranlund for helping spark the idea for this research and for providing feedback in the early stages. The paper has benefited from comments by John Heywood, Juan Moreno-Cruz, participants at the 2017 meeting of the Southern Economic Association, seminar participants at Northern Illinois University and two anonymous referees.

A. Appendix

A.1. Proof of Proposition 1

The Nash equilibrium tax rates are the simultaneous solution to the signatory (20) and non-signatory (21) best-response functions. Substituting (21) into (20) results in

$$\begin{aligned}
 p_s &= \frac{s \left[b\gamma - \beta \left[(m-s) \left[\frac{b\gamma - \beta [sp_s - cm]}{\beta(m-s) + \gamma} \right] - mc \right] \right]}{s^2\beta + \gamma} \\
 p_s &= \frac{s b\gamma - s\beta(m-s) \left[\frac{b\gamma - \beta [sp_s - cm]}{\beta(m-s) + \gamma} \right] + s\beta mc}{s^2\beta + \gamma} \\
 p_s &= \frac{\left[s b\gamma^2 + s\beta b\gamma(m-s) - s\beta b\gamma(m-s) \right. \\
 &\quad \left. + s\beta^2(m-s)[sp_s - cm] + s\beta^2 mc(m-s) + s\beta\gamma mc \right]}{[s^2\beta + \gamma][\beta(m-s) + \gamma]} \\
 p_s &= \frac{s b\gamma^2 + s\beta^2(m-s)[sp_s - cm] + s\beta^2 mc(m-s) + s\beta\gamma mc}{[s^2\beta + \gamma][\beta(m-s) + \gamma]} \\
 p_s &= \frac{s b\gamma^2 + s^2\beta^2(m-s)p_s + s\beta\gamma mc}{[s^2\beta + \gamma][\beta(m-s) + \gamma]} \\
 p_s &= \frac{s b\gamma^2 + s^2\beta^2(m-s)p_s + s\beta\gamma mc}{[s^2\beta^2(m-s) + \beta(m-s)\gamma + s^2\beta\gamma + \gamma^2]}.
 \end{aligned}$$

Then multiplying both sides by the denominator in brackets yields

$$\begin{aligned}
 p_s [s^2\beta^2(m-s) + \beta(m-s)\gamma + s^2\beta\gamma + \gamma^2] &= s b\gamma^2 + s^2\beta^2(m-s)p_s + s\beta\gamma mc \\
 p_s [\beta(m-s)\gamma + s^2\beta\gamma + \gamma^2] &= s b\gamma^2 + s\beta\gamma mc \\
 p_s [\beta(m-s) + s^2\beta + \gamma] &= s b\gamma + s\beta mc \\
 p_s [\beta(s^2 + m-s) + \gamma] &= s [b\gamma + \beta mc] \\
 p_s &= \frac{s [b\gamma + \beta mc]}{\beta(s^2 + m-s) + \gamma}. \tag{A.1}
 \end{aligned}$$

Now we can substitute the signatory emissions tax (A.1) into the non-signatory best-response (21) to obtain the non-signatories' emissions tax.

$$\begin{aligned}
 p_t &= \frac{b\gamma - \beta \left[s \left(\frac{s[b\gamma + \beta mc]}{\beta(s^2 + m-s) + \gamma} \right) - cm \right]}{\beta(m-s) + \gamma} \\
 p_t &= \frac{b\gamma^2 + b\gamma\beta(s^2 + m-s) - \beta s^2 [b\gamma + \beta mc] + cm\beta^2(s^2 + m-s) + cm\beta\gamma}{[\beta(m-s) + \gamma][\beta(s^2 + m-s) + \gamma]} \\
 p_t &= \frac{b\gamma^2 + b\gamma\beta(m-s) - \beta s^2 [\beta mc] + cm\beta^2(s^2 + m-s) + cm\beta\gamma}{[\beta(m-s) + \gamma][\beta(s^2 + m-s) + \gamma]} \\
 p_t &= \frac{b\gamma^2 + b\gamma\beta(m-s) + cm\beta^2(m-s) + cm\beta\gamma}{[\beta(m-s) + \gamma][\beta(s^2 + m-s) + \gamma]} \\
 p_t &= \frac{\beta(m-s)[cm\beta + b\gamma] + b\gamma^2 + cm\beta\gamma}{[\beta(m-s) + \gamma][\beta(s^2 + m-s) + \gamma]}
 \end{aligned}$$

$$\begin{aligned}
p_t &= \frac{\beta(m-s)[cm\beta + b\gamma] + \gamma[cm\beta + b\gamma]}{[\beta(m-s) + \gamma][\beta(s^2 + m - s) + \gamma]} \\
p_t &= \frac{[\beta(m-s) + \gamma][cm\beta + b\gamma]}{[\beta(m-s) + \gamma][\beta(s^2 + m - s) + \gamma]} \\
p_t &= \frac{\beta cm + b\gamma}{\beta(s^2 + m - s) + \gamma}.
\end{aligned} \tag{A.2}$$

A.2. Derivation of signatory and non-signatory abatement

Substituting p_s from [Proposition 1](#) into Eq. (5) and $p = p_s = \frac{s[b\gamma + \beta mc]}{\beta(s^2 + m - s) + \gamma}$ results in

$$\begin{aligned}
\gamma x_s &= p_s - c \\
\gamma x_s &= \frac{s[b\gamma + \beta mc]}{\beta(s^2 + m - s) + \gamma} - c \\
\gamma x_s &= \frac{s[b\gamma + \beta mc] - c\beta(s^2 + m - s) - c\gamma}{\beta(s^2 + m - s) + \gamma} \\
x_s &= \frac{\gamma(sb - c) + s\beta mc - c\beta(s^2 + m - s)}{\beta\gamma(s^2 + m - s) + \gamma^2} \\
x_s &= \frac{\gamma(sb - c) + c\beta[sm - s^2 - m + s]}{\beta\gamma(s^2 + m - s) + \gamma^2} \\
x_s &= \frac{\gamma(sb - c) + c\beta[s(m - s) - (m - s)]}{\beta\gamma(s^2 + m - s) + \gamma^2} \\
x_s &= \frac{\gamma(sb - c) + c\beta[(m - s)(s - 1)]}{\beta\gamma(s^2 + m - s) + \gamma^2},
\end{aligned}$$

which is Eq. (19).

Similarly, substituting p_t from [Proposition 1](#) into Eq. (5) while recognizing that $p_t = \frac{\beta cm + b\gamma}{\beta(s^2 + m - s) + \gamma}$ results in

$$\begin{aligned}
\gamma x_t &= \frac{\beta cm + b\gamma}{\beta(s^2 + m - s) + \gamma} - c \\
\gamma x_t &= \frac{\beta cm + b\gamma - c\beta(s^2 + m - s) - c\gamma}{\beta(s^2 + m - s) + \gamma} \\
x_t &= \frac{\gamma(b - c) - c\beta(s^2 - s)}{\beta\gamma(s^2 + m - s) + \gamma^2} \\
x_t &= \frac{\gamma(b - c) - c\beta s(s - 1)}{\beta\gamma(s^2 + m - s) + \gamma^2},
\end{aligned}$$

which is Eq. (20).

A.3. Proof of [Proposition 3](#)

The Proof proceeds as follows (i) we show that an IEA of size $s = 2$ can be stable under certain conditions, and (ii) an IEA of size $s \geq 3$ cannot be stable under any condition. To accomplish (ii), we show that the IS condition for an IEA of size $s \geq 3$ cannot change sign for any positive parameters. This means that if an IEA is not internally stable for any $s \geq 3$ then it is unstable for all $s \geq 3$. (iii) We then show that the grand coalition $s = m \geq 3$ is unstable, hence all IEAs with more than two members are unstable.

(i) To determine if a two member IEA can be internally stable we can evaluate m^* in Eq. (28) at $s = 2$ to obtain

$$m^*(s = 2, \beta, \gamma) = \frac{2\beta - \gamma + 2\sqrt{(\beta + \gamma)(4\beta + \gamma)}}{3\beta}. \tag{A.3}$$

Given that the number of signatories s must be less than or equal to the number of players m an internally stable IEA is possible if $m^*(s = 2, \beta, \gamma) \geq 2$, which yields

$$m^*(s = 2, \beta, \gamma) = \frac{2\beta - \gamma + 2\sqrt{(\beta + \gamma)(4\beta + \gamma)}}{3\beta} \geq 2$$

$$3\gamma(4\beta + \gamma) \geq 0. \quad (\text{A.4})$$

(ii) We will now demonstrate that no stable IEA of size $s > 2$ can exist. In order to demonstrate this, we will first show that $m^*(s, \beta, \gamma) < 1$ for all IEAs with $s > 2$. That is, we can show that

$$m^*(s, \beta, \gamma) = \frac{-\beta s(s^2 - 2s - 1) - \gamma(s - 1) + 2\sqrt{(\beta + \gamma)(\beta s^2 + \gamma)}}{\beta(s + 1)} < 1,$$

for all $s > 2$.

$$\begin{aligned} -\beta s(s^2 - 2s - 1) - \gamma(s - 1) + 2\sqrt{(\beta + \gamma)(\beta s^2 + \gamma)} &< \beta(s + 1) \\ 2\sqrt{(\beta + \gamma)(\beta s^2 + \gamma)} &< \beta[s(s^2 - 2s - 1) + s + 1] + \gamma(s - 1) \\ 2\sqrt{(\beta + \gamma)(\beta s^2 + \gamma)} &< \beta[s(s^2 - 2s) + 1] + \gamma(s - 1) \\ \beta^2[s^2(s - 2) + 1]^2 + 2\beta\gamma(s - 1)[s^2(s - 2) + 1] + \gamma^2(s - 1)^2 &> 4(\beta + \gamma)(\beta s^2 + \gamma) \\ \beta^2[s^2(s - 2) + 1]^2 + 2\beta\gamma(s - 1)[s^2(s - 2) + 1] + \gamma^2(s - 1)^2 &> 4[\beta^2 s^2 + \beta\gamma + \beta\gamma s^2 + \gamma^2] \end{aligned}$$

subtracting the right-hand-side from both sides and gathering terms yields

$$\begin{aligned} \beta^2([s^2(s - 2) + 1]^2 - 4s^2) + 2\beta\gamma[(s - 1)[s^2(s - 2) + 1] - 2s^2 - 2] + \gamma^2[(s - 1)^2 - 4] &> 0 \\ \beta^2[s^4(s - 2)^2 + 2s^2(s - 4) + 1] + 2\beta\gamma[s^3(s - 2) + s - s^3 - 3] + \gamma^2[s^2 - 2s - 3] &> 0 \\ \beta^2[s^4(s - 2)^2 + 2s^2(s - 4) + 1] + 2\beta\gamma[s^3(s - 3) + s - 3] + \gamma^2[(s + 1)(s - 3)] &> 0 \\ \beta^2[s^4(s - 2)^2 + 2s^2(s - 4) + 1] + 2\beta\gamma(s^3 + 1)(s - 3) + \gamma^2(s + 1)(s - 3) &> 0 \end{aligned}$$

which is satisfied for all $s > 2$. Note if $s = 3$ then last line above reduces to $64\beta^2 > 0$ and all three terms are clearly positive for $s > 3$. Therefore $m^*(s, \beta, \gamma) < 1$ for all $s > 2$ meaning the internal stability condition cannot change sign in the relevant range where $s \leq m$.

(iii) Suppose $s = m = 3$. The IS condition given the grand coalition (i.e., $s = m$) simplifies to

$$\pi_i^s(s = m = 3) - \pi_i^t(m = 1) = \frac{-8\beta(b^2\gamma^2 + 6bc\beta\gamma + 9c^2\beta^2)}{\gamma(5\beta + \gamma)^2(9\beta + \gamma)} < 0. \quad (\text{A.5})$$

Since the IS condition is violated (negative), no three member IEA will be stable even if there are only three countries. Repeating this procedure for larger values of $m > 3$ shows that the IS condition for the grand coalition is strictly negative. Since the grand coalition is not internally stable, and $m^*(s, \beta, \gamma) < 1$ implies the IS condition cannot change sign for $s \geq 3$, coalitions of size $s \geq 3$ are unstable.

A.4. Proof of Proposition 4

When countries are identical, each has a payoff of

$$\pi_i = bX - \frac{\beta}{2}X^2 - cx_i - \frac{\gamma}{2}x_i^2. \quad (\text{A.6})$$

Signatories maximize the aggregate coalition payoff, $v(s) = \sum_{i=1}^s \pi_i$, which is

$$v(s) = \sum_{i=1}^s \pi_i = \sum_{i=1}^s \left[bX - \frac{\beta}{2}X^2 - cx_i - \frac{\gamma}{2}x_i^2 \right]. \quad (\text{A.7})$$

Rewriting (A.7) in terms of abatement by signatories x_i^s and nonsignatories x_j^t yields

$$v(s) = \sum_{i=1}^s \left[b \left(\sum_{i=1}^s x_i^s + \sum_{j=s+1}^m x_j^t \right) - \frac{\beta}{2} \left(\sum_{i=1}^s x_i^s + \sum_{j=s+1}^m x_j^t \right) - c x_i^s - \frac{\gamma}{2} (x_i^s)^2 \right] \quad (\text{A.8})$$

and the first-order condition is

$$\sum_{i=1}^s b - \beta \sum_{i=1}^s \left(\sum_{i=1}^s x_i^s + \sum_{j=s+1}^m x_j^t \right) - c - \gamma x_i^s = 0. \quad (\text{A.9})$$

Since each signatory has an identical first-order condition $\sum_{i=1}^s = s$ and $x_i^s = x_s$ for all $i = 1, \dots, s$ this becomes

$$sb - c - \beta s^2 x_s - \beta s \sum_{j=s+1}^m x_j^t - \gamma x_s = 0. \quad (\text{A.10})$$

Solving for x_s yields the signatory best-response function

$$x_s^{br} = \frac{sb - c - \beta s \sum_{j=s+1}^m x_j^t}{\beta s^2 + \gamma}. \quad (\text{A.11})$$

Meanwhile, non-signatories choose abatement to maximize individual payoff in (A.6). The first-order condition for country k is

$$b - \beta \left(\sum_{i=1}^s x_i^s + \sum_{j=s+1}^m x_j^t \right) - c - \gamma x_k^t = 0. \quad (\text{A.12})$$

Given the symmetric FOC for any two non-signatories implies $x_j^t = x_k^t = x_t$ and $\sum_{j=s+1}^m = (m - s)$. This becomes

$$b - \beta \sum_{i=1}^s x_i^s - \beta(m - s)x_t - c - \gamma x_t = 0, \quad (\text{A.13})$$

which results in the non-signatory best-response function

$$x_t^{br} = \frac{b - c - \beta \sum_{i=1}^s x_i^s}{\gamma + \beta(m - s)}. \quad (\text{A.14})$$

The Nash equilibrium is the simultaneous solution to (A.11) and (A.14). Substituting (A.14) into (A.11) and simplifying results in

$$\begin{aligned} x_s^{br} &= \frac{sb - c - \beta s(m - s)x_t^{br}}{\beta s^2 + \gamma} \\ x_s [\beta s^2 + \gamma] &= sb - c - \beta s(m - s) \left[\frac{b - c - \beta s x_s}{\gamma + \beta(m - s)} \right] \\ x_s [\beta s^2 + \gamma] [\gamma + \beta(m - s)] &= (sb - c) [\gamma + \beta(m - s)] - \beta s(m - s) [b - c - \beta s x_s] \\ x_s [\beta \gamma (s^2 + m - s) + \gamma^2 + \beta^2 s^2(m - s)] &= \gamma (sb - c) + \beta(m - s) [(sb - c) - sb + sc] + \beta^2 s^2(m - s)x_s \\ x_s [\beta \gamma (s^2 + m - s) + \gamma^2] &= \gamma (sb - c) + \beta(m - s) [c(s - 1)] \\ x_s^{ne} &= \frac{\gamma (sb - c) + c\beta(m - s)(s - 1)}{\beta \gamma (s^2 + m - s) + \gamma^2} \end{aligned} \quad (\text{A.15})$$

Nash equilibrium signatory abatement from payoff maximization with respect to quantity in (A.15) results in the same solution as the price maximization problem in (22). Similarly, substituting (A.11) into (A.14) results in

$$\begin{aligned} x_t^{br} &= \frac{b - c - \beta s x_s^{br}}{\gamma + \beta(m - s)} \\ x_t [\gamma + \beta(m - s)] &= b - c - \beta s \left[\frac{sb - c - \beta s(m - s)x_t}{\beta s^2 + \gamma} \right] \\ x_t [\gamma + \beta(m - s)] [\beta s^2 + \gamma] &= (b - c) [\beta s^2 + \gamma] - \beta s [sb - c - \beta s(m - s)x_t] \\ x_t [\beta \gamma s^2 + \beta \gamma(m - s) + \beta^2 s^2(m - s) + \gamma^2] &= \gamma (b - c) + \beta s^2 (b - c) - \beta s^2 b + c\beta s + \beta^2 s^2(m - s)x_t \end{aligned}$$

$$x_t [\beta\gamma (s^2 + m - s) + \gamma^2] = \gamma (b - c) - c\beta s^2 + c\beta s$$

$$x_t^{ne} = \frac{\gamma (b - c) - c\beta s(s - 1)}{\beta\gamma (s^2 + m - s) + \gamma^2} \quad (\text{A.16})$$

which is the same solution as the price maximization problem in (23). Recognizing that the emissions tax p is equal to the marginal abatement cost of the last unit $p_i = c + \gamma x_i$ results in $x_i = \frac{p_i - c}{\gamma}$. Hence the stability condition in (27) is identical for the price and quantity maximization problems.

References

- Aldy, J.E., 2015. Pricing climate risk mitigation. *Nat. Clim. Change* 5, 396–398.
- Barrett, S., 1994. Self-enforcing international environmental agreements. *Oxf. Econ. Pap.* 46, 878–894.
- Barrett, S., 2005. The theory of international environmental agreements. In: Maler, K.G., Vincent, J.R. (Eds.), *Handbook of Environmental Economics*, vol. 3. Elsevier.
- Barrett, S., 2006. Climate treaties and breakthrough technologies. *Am. Econ. Rev.* 96 (2), 22–25.
- Battaglini, M., Harstad, B., 2016. Participation and duration of environmental agreements. *J. Polit. Econ.* 124 (1), 160–204.
- Carraro, C., Siniscalco, D., 1993. Strategies for the international protection of the environment. *J. Publ. Econ.* 52 (3), 309–328.
- Cooper, R.N., 2010. The case for charges on greenhouse gas emissions. In: Aldy, J., Stavins, R. (Eds.), *Post-kyoto International Climate Policy: Architectures for Agreement*. Cambridge University Press.
- Cramton, P., Ockenfels, A., Stoft, S., 2015. An international carbon price commitment promotes cooperation. *Econ. Energy Environ. Pol.* 4 (2), 51–64.
- d'Aspremont, C.A., Jacquemin, A., Gabszewicz, J.J., Weymark, J.A., 1983. On the stability of collusive price leadership. *Can. J. Econ.* 16 (1), 17–25.
- de Zeeuw, A., 2015. International environmental agreements. *Annu. Rev. Resour. Econ.* 7, 151–168.
- Diamantoudi, E., Sartzetakis, E.S., 2006. Stable international environmental agreements: an analytical approach. *J. Publ. Econ. Theor.* 8 (2), 247–263.
- Dixit, A., Olson, M., 2000. Does voluntary participation undermine the Coase theorem? *J. Publ. Econ.* 75, 309–335.
- Edenhofer, O., Jakob, M., Creutzig, F., Flachsland, C., Fuss, S., Kowarsch, M., Lessmann, K., Mattau, L., Siegmeier, J., Steckel, J.C., 2015. Closing the emission price gap. *Global Environ. Change* 31, 132–143.
- Eichner, T., Pethig, R., 2013. Self enforcing environmental agreements and international trade. *J. Publ. Econ.* 102, 37–50.
- Ellingsen, T., Paltseva, E., 2016. Confining the Coase theorem: contracting, ownership and free riding. *Rev. Econ. Stud.* 83, 547–586.
- Endres, A., 1997. Negotiating a climate convention: the role of prices and quantities. *Int. Rev. Law Econ.* 17, 147–156.
- Feldstein, M.S., Halstead, T., Mankiw, N.G., February 8 2017. A Conservative Case for Climate Action. *The New York Times*.
- Finus, M., Maus, S., 2008. Modesty may pay. *J. Publ. Econ. Theor.* 10 (5), 801–826.
- Finus, M., Rundshagen, B., 2003. Endogenous coalition formation in global pollution control: a partition function approach. In: Carraro, C. (Ed.), *The Endogenous Formation of Economic Coalitions*. Edward Elgar, pp. 199–243 (chapter 6).
- Gelves, J.A., McGinty, M., 2016. International environmental agreements with consistent conjectures. *J. Environ. Econ. Manag.* 78, 67–84.
- Goeschl, T., Perino, G., 2017. The climate policy hold-up: green technologies, intellectual property rights, and the abatement incentives of international agreements. *Scand. J. Econ.* 119, 709–732.
- Hoel, M., 1991. Efficient international agreements for reducing emissions of CO₂. *Energy J.* 12 (2), 93–107.
- Hoel, M., 1992. Carbon taxes: an international tax or harmonized domestic taxes? *Eur. Econ. Rev.* 36, 400–406.
- Hoel, M., Karp, L., 2002. Taxes vs. quotas for a stock pollutant. *Resour. Energy Econ.* 24, 367–384.
- Kolstad, C., 2007. Systematic uncertainty in self enforcing international environmental agreements. *J. Environ. Econ. Manag.* 53, 68–78.
- McEvoy, D.M., Stranlund, J.K., 2009. Self-enforcing international environmental agreements with costly monitoring for compliance. *Environ. Resour. Econ.* 42, 491–508.
- McGinty, M., 2007. International environmental agreements among asymmetric nations. *Oxf. Econ. Pap.* 59 (1), 45–62.
- Nordhaus, W.D., 1982. How fast should we graze the global commons? *Am. Econ. Rev.* 72 (2), 242–246.
- Nordhaus, W.D., 2006. After kyoto: alternative mechanisms to control global warming. *Am. Econ. Rev.* 96 (2), 31–34.
- Nordhaus, W.D., 2007. To tax or not to tax: alternative approaches to slowing global warming. *Rev. Environ. Econ. Pol.* 1 (1), 26–44.
- Nordhaus, W.D., 2015. Climate clubs: overcoming free riding in international climate policy. *Am. Econ. Rev.* 105 (4), 1339–1370.
- Pearce, D., 1992. The role of carbon taxes in adjusting to global warming. *Econ. J.* 101 (407), 938–948.
- Rubio, S.J., Ulph, A., 2007. An infinite horizon model of dynamic membership of international environmental agreements. *J. Environ. Econ. Manag.* 54 (3), 296–310.
- Ulph, A., 2004. Stable international agreements with a stock pollutant. *J. Risk Uncertain.* 29 (1), 53–73.
- van der Ploeg, F., De Zeeuw, A.J., 1992. International aspects of pollution control. *Environ. Resour. Econ.* 2, 117–139.
- Weitzman, M., 1974. Prices vs. quantities. *Rev. Econ. Stud.* 41 (4), 477–491.
- Weitzman, M., 2014. Can negotiating a uniform carbon price help to internalize the global warming externality? *J. Assoc. Environ. Resour. Econ.* 1 (1/2), 29–49.
- Weitzman, M., 2016. Voting on prices vs. quantities in a world climate assembly. *Res. Econ.*, <https://doi.org/10.1016/j.rie.2016.10.004>.