Incentives to Join International Environmental Agreements with Permit Trading and Safety Valves*

Larry Karp†

April 2, 2010

Abstract

Conditional on the level of membership of an International Environmental Agreement, trade in emissions allowances reduces aggregate abatement costs and promotes efficiency. Using a safety valve (a price ceiling) together with trade leads to a further increase in efficiency, conditional on membership. However, both of these policies reduce the extent to which a potential signatory to an International Environmental Agreement is "pivotal", thereby reducing incentives to join the agreement. The safety valve, despite increasing efficiency conditional on membership, causes a more pronounced reduction in the incentive to participate. A non-standard, "high-powered" safety valve can promote membership and achieve ex post efficiency.

Keywords: Kyoto Protocol, emission permit trading, safety valve, cost uncertainty, participation game, International Environmental Agreement.

JEL classification numbers C72, H4, Q54

*I thank Jinhua Zhao for extensive conversations on the background for this paper and the Ragnar Frisch Centre for Economic Research for financial support.

†Department of Agricultural and Resource Economics, University of California, Berkeley, and the Ragnar Frisch Center for Economic Research, email: karp@berkeley.edu

1 Introduction

The design of the successor to the Kyoto Protocol should encourage countries to join the agreement ("ex ante efficiency"), and it should lead to a high level of welfare for signatories conditional on membership ("ex post efficiency"). Allowing signatories to trade emissions permits amongst themselves, thereby achieving an efficient allocation of abatement effort, is part of the Kyoto Protocol and will likely be part of the next agreement. A price ceiling together with trade in permits – a combination known as a "safety valve" – provides an extra instrument that increases ex post efficiency, and also protects signatories against unexpectedly high abatement costs. Thus, both policies promote ex post efficiency, and the safety valve does so to a greater extent than trade alone. However, both policies undermine incentives to participate, and the safety valve does so to a greater extent than trade alone.

Arguably the greatest impediment to a successful climate agreement is the difficulty of persuading nations to provide the global public good of abatement, i.e. getting nations to join and abide by an agreement that reduces emissions. The problem of achieving abatement efficiently may be a less serious design problem. If this judgement is correct, trade in permits, with or without a price ceiling, may not help solve the problem of climate change. However, a slightly more complex safety valve can promote both ex ante and ex post efficiency.

A large literature models international environmental agreements (IEAs) as a non-cooperative Nash equilibrium of a participation game (Hoel 1992, Carraro and Siniscalco 1993, Barrett 1994, 2003, 2006, Bloch 1997, Dixit 2000, Finus 2001, and Batabyal 2002).[1] A "standard model" consists of a two-stage game. In the first stage (the participation game) countries decide whether to join an IEA; in the second stage (the abatement game), the IEA takes an action from an exogenously given menu to maximize members' expected welfare.

In a special case of the standard model where costs and benefits enter payoffs linearly, members of the IEA decide to abate at capacity or not at all, and non-signatories' dominant strategy is to not abate. A potential member to the IEA can affect the abatement actions (but not the participation decisions) only of other members, so the potential member's incentive to join depends on the leverage it exerts over those members. A member is pivotal if, by joining the IEA, it has a substantial effect on the behavior of other members. In this model, the unique

[1] The Nash assumption means that when deciding to join, a country takes other countries' membership decisions as given, and therefore does not try to influence those decisions. In more sophisticated models of coalition formation, an agent has beliefs about how its provisional decision to join or leave a coalition would affect other agents' participation decisions (Chwe 1994), (Mariotti 1997), (Xue 1998), and (Ray and Vohra 2001). Diamantoudi and Sartzetakis (2002), Eyckmans (2001), and de Zeeuw (2008) use models of this sort to study IEAs.

A growing literature uses experimental methods to assess the plausibility of different equilibria in participation games (Burger and Kolstad 2009), (Kosfeld, Okada, and Riedl 2009), (Dannenberg, Lange, and Sturm 2009).

non-trivial pure strategy Nash equilibrium contains the minimal number of members for which abatement at capacity is optimal. The major result from the standard model is that a decrease in abatement costs, or more generally an increase in the potential gain from cooperation, weakly decreases the size of this IEA and potentially decreases global welfare.

Trade in emissions permits weakly decreases expected abatement costs; the ability to use a price ceiling together with trade further increases members' welfare, conditional on the level of membership. In view of the major result from the standard model, described above, it might appear not surprising that emissions trade, with or without the price ceiling, weakly lowers equilibrium membership and potentially lowers welfare. However, the effects of trade and the price ceiling are much more complicated than are the effects of an exogenous decrease in abatement costs. With trade, the addition of a new member can lead to a greater increase in average abatement of other members, compared to the no-trade case; in this respect, trade increases the leverage of a potential member, a fact that tends to promote greater equilibrium membership. However, with emissions trade, the defection of a member leads to a smaller equilibrium reduction in remaining members' average abatement, relative to the no-trade case. This fact tends to undermine the participation incentive, because it reduces the credible threat of the IEA to make a drastic reduction in abatement, following a defection. These comparisons are even more striking in the case of trade with a price ceiling. Because trade, with or without the price ceiling, creates these two opposing effects on membership incentives, the equilibrium effect of trade appears to be ambiguous. A formal model helps to resolve this apparent ambiguity.

A price ceiling that is optimal conditional on membership gets the direction of incentives right, because a larger number of members increases the equilibrium price ceiling, thereby increasing the average abatement of each member. However, this ex post optimal price ceiling is not sufficiently "high-powered" to offset the fact that it undermines the IEA's credible threat to drastically curtail abatement following a defection. A slightly more sophisticated design of the price ceiling can overcome this problem, and induce a high level of participation and a high degree of ex post efficiency.

2 Preliminaries

There are N ex ante identical countries with constant abatement costs up to a maximum level, normalized to 1. Abatement is a global public good, and the marginal expected benefit of abatement, for each country, is a constant, normalized to 1. Country i's constant abatement cost, up to its capacity, is θ_i.

In the first stage of a two-stage game, nations decide whether to join the IEA and in the second stage the IEA chooses members' required abatement level and/or the price ceiling. At the participation stage (stage 1), all nations are identical. It is common knowledge that costs are independently and identically distributed with support $[\theta_L, \theta_H]$, with $\theta_L > 1$ and $E\theta \equiv \bar{\theta}$. At the participation stage, nations do not know their cost realization. At the abatement stage (stage 2), firms know their own costs, but this information is not verifiable, so it is not possible to write a contract on the cost realization. At the abatement stage, non-signatories have a dominant strategy, not to abate, because $\theta_L > 1$. The agreement that signatories signed at the participation stage constrains their behavior at the abatement stage.

I consider three types of simple agreement. In the "standard case", the IEA chooses the level of abatement conditional on membership. Because costs are non-verifiable, the IEA instructs each signatory to abate at the same level. The second type of simple agreement uses cap and trade; the IEA assigns each signatory a required abatement level, but allows the signatories to satisfy this requirement by trading emissions permits with other signatories (but not with non-signatories). The third type of simple agreement uses cap and trade together with a safety valve; the IEA assigns a required abatement level and allows trade, and in addition it sets a constant price ceiling. Countries are allowed to buy additional emissions permits at this price ceiling. The type of agreement, "no-trade", "trade" or "safety valve", is exogenous: the purpose of the model is to determine how the type of agreement affects equilibrium behavior.

I emphasize the time line under which the IEA chooses the policy level in order to maximize members' expected welfare conditional on the level of membership. This equilibrium is "renegotiation proof". Under the alternative timing, where the IEA makes its decision before participation decisions, the equilibrium IEA has no members with all three simple agreements. However, use of the more sophisticated price ceiling does require the IEA to commit to a policy prior to countries' participation decisions (Section 4).

The assumption that countries are identical at the participation stage does not describe the real world, but it is useful because it emphasizes the role of trade in achieving efficiency when cost differences are not verifiable.[2] If costs could be verified by an international institution, the IEA could make the abatement allocation conditional on those costs, eliminating one reason for trade.

There is substantial variation in estimated economic costs of reducing GHG emissions (Aldy, Krupnick, Newell, Parry, and Pizer 2008), (Fischer and Morgenstern 2006). Some estimates, particularly those commissioned by industry groups, find very high costs. Other estimates assume that win-win policies abound, leading to low abatement costs. Actual abatement costs are likely to be highly positively correlated, a feature that the model does not capture. For example, the successful development of carbon capture and storage technologies will re-

[2]Barrett (2001), Fujita (2006), Kolstad and Ulph (2008), and McGinty (2007) study IEA models where countries are different at the participation stage.

duce costs everywhere. An extension of the current model that includes a common cost shock in addition to the country-specific shock would make the model more descriptive, but would not add to the insight it provides about trade.

The next climate agreement is likely to restrict emissions over a short period of time, perhaps a decade. During that period we are likely to learn more about the cost of abatement than about climate damages. This assumption motivates the emphasis on cost uncertainty, to the neglect of damage uncertainty.

A review: the IEA without trade In the "standard model", the IEA chooses the level of abatement that maximizes signatories' expected welfare and does not permit trade. Since costs are non-verifiable, the IEA assigns the same required level of abatement to all signatories. It conditions the abatement level on M, the number of signatories. Define $h(x)$ as the smallest integer not less than x; $h(\cdot)$ is the "ceiling function". (A mnemonic:"h" for "heaven".) In the absence of trade, an IEA with M members sets per member abatement equal to

$$a(M) = \begin{cases} 1 \text{ if } M \geq h\left(\bar{\theta}\right) \\ 0 \text{ if } M < h\left(\bar{\theta}\right), \end{cases}$$

and the unique non-trivial Nash equilibrium membership (the "no-trade IEA") is

$$M = h\left(\bar{\theta}\right). \tag{1}$$

There are trivial equilibria with $M < h\left(\bar{\theta}\right)$ in which the IEA instructs its members not to abate, but I ignore those. The (non-trivial) equilibrium expected payoff for a signatory, $\pi^{s,\text{no trade}}$, satisfies

$$1 \geq \pi^{s,\text{no trade}} \equiv h\left(\bar{\theta}\right) - \bar{\theta} \geq 0, \tag{2}$$

and the payoff for a non-signatory equals $h\left(\bar{\theta}\right)$. In a Nash equilibrium at the participation stage, a signatory believes that if it defects, the resulting IEA will set $a(\left(h\left(\bar{\theta}\right) - 1\right) = 0$, so a nation that defects would obtain a payoff of 0. In equilibrium, no member wants to defect and no non-member wants to join.

3 Trade with and without a price ceiling

This section analyzes the simple agreements that use cap and trade, first without and then with a price ceiling. In both cases countries are price takers in the permit market. My main interest is in the time-line where the IEA conditions its decision (the abatement target and/or price ceiling) on the number of members, M. However, the reversed timing, where the IEA makes its decision prior to nations' participation decisions, is useful for understanding the more sophisticated price ceiling (Section 4). The following is an obvious implication of the model:

Remark 1 *If the IEA chooses the level of abatement before the participation decisions, a potential signatory's participation decision has no effect on signatories' abatement level, regardless of whether the IEA allows trade at the second stage. Therefore, for this time line, no nation has an incentive to join, and the unique Nash equilibrium to the participation game is* $M = 0$.

I show below that if the IEA chooses a price ceiling prior to the participation decisions, the unique Nash equilibrium to the participation game is again $M = 0$ (Proposition 4). There, however, the conclusion is not obvious, because an additional member potentially changes the probability that a fixed price ceiling is binding. In this case, an additional member potentially changes the expected level of abatement of existing members, and thus nations may have an incentive to join. For this reason, a different kind of argument must be used to establish that in equilibrium no nation chooses to join an IEA with a predetermined price ceiling.

Under the standard time line, where the IEA conditions its decision on membership, the trade regimes with or without the price ceiling both create an incentive for non-members to participate, because under both regimes a non-member's participation (typically) increases expected abatement of existing members. However, both trade regimes also weaken the credible penalty to defection, because a member knows that its defection leads to a reduction, but not a cessation, of abatement. The second effect is more powerful, so on balance both trade regimes (typically) reduce equilibrium participation, relative to the standard no-trade regime.

3.1 The IEA with cap and trade

The optimal cap for the IEA depends on the number of members and the distribution of costs. Denote the i'th order statistic for costs when there are M signatories, and the expectation of this order statistic as, respectively,

$$\theta_{(i,M)} = \theta_{(i)} \mid M, \text{ and } \bar{\theta}_{(i,M)} = E\theta_{(i)} \mid M.$$

Suppose that the IEA requires each signatory to abate at level $\bar{a} = \frac{k}{M} < 1$; $\bar{a}$ is the average level of abatement per signatory and k is the aggregate level of abatement. Signatories are allowed to achieve their target by trading amongst themselves, but not with non-members. If k is an integer, the k members with the lowest costs abate at capacity, and the remaining members buy permits from them. Signatories are price takers in the permit market, so trade is efficient.

The step function in Figure 1 shows the expected marginal cost per member as average abatement, $\bar{a}$, ranges from 0 to 1. For example, if $M = 10$ and $k = 5.4$, the signatory with the 6'th highest cost abates partially, so the expected marginal cost of an increase in $\bar{a}$ is $\bar{\theta}_{(6,10)}$.

A marginal increase in $\bar{a}$ causes a marginal increase in total abatement equal to M, so the marginal benefit per member equals the constant M, as Figure 1 shows. When the horizontal

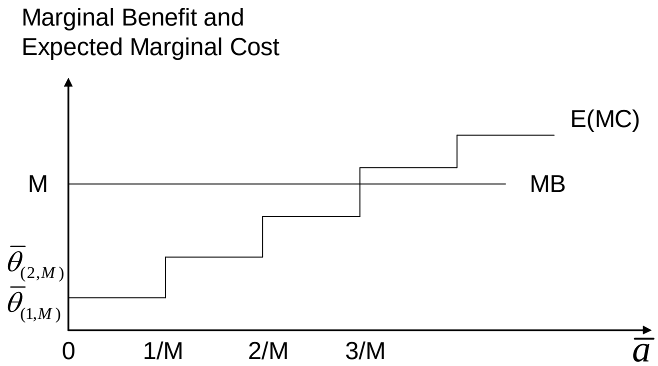

Figure 1: Per-member marginal benefit and expected marginal costs with trade

line at M intersects a vertical part of the step function, this intersection is the unique optimal (for IEA members) level of $\bar{a}$, so k is an integer If the horizontal line intersects the marginal cost curve on a flat portion, a range of values is optimal. The tie-breaking assumption is that the IEA chooses the maximum optimal level of $\bar{a}$, so once again k is an integer.

With this tie-breaking assumption, the equilibrium level of per signatory abatement is $\bar{a} = \frac{k}{M}$, where k is the integer that solves

$$k = \max_{1 \leq i \leq M} i \text{ subject to } \bar{\theta}_{(i,M)} \leq M. \tag{3}$$

Trade occurs if and only if $k < M$ in equilibrium.

In general, trade can increase or decrease the equilibrium level of abatement for given M. In the absence of trade, and given non-verifiable costs, the IEA requires every country to abate at the same level. Up to the capacity constraint, the expected marginal cost of increasing abatement in the absence of trade is simply $\bar{\theta}$, a constant. With trade, the expected marginal cost curve is the step function in Figure 1, with the first steps lower than $\bar{\theta}$ and latter steps higher than $\bar{\theta}$. The per-member marginal benefit of abatement is M, regardless of whether there is trade. Therefore, depending on the level of M, the optimal level of abatement under trade can be either higher or lower than the optimal level in the absence of trade. That is, even though trade reduces the average aggregate expected cost of achieving any level of abatement short of capacity, trade can increase the expected marginal costs of abatement, and therefore can decrease the level of abatement that the IEA chooses.

In the IEA game, M is endogenous. After discussing some features of trade with a general distribution of costs, I focus on the uniform distribution in order to obtain clear results.

Denote signatory j's cost of achieving the target level of abatement under trade as c_j. Some of these costs may be negative, e.g. if nation j sells permits, and the permit price is sufficiently high and the required level of abatement is sufficiently low. Because trade is efficient, the sum of signatories' costs is

$$\sum_{j=1}^{M} c_j = \sum_{i=1}^{k} \theta_{(i,M)},$$

so the expectation of total abatement costs is

$$\sum_{i=1}^{k} \bar{\theta}_{(i,M)}. \tag{4}$$

Since each signatory has the same probability of any ranking, all signatories have the same expected cost:

$$\frac{\sum_{i=1}^{k} \bar{\theta}_{(i,M)}}{M}.$$

Consider the effect of allowing trade in an IEA that has formed under the belief that trade will not be allowed, i.e. where $M = h\left(\bar{\theta}\right)$. A necessary and sufficient condition for trade to occur when it is allowed is that k, defined in equation (3), is less than $M = h\left(\bar{\theta}\right)$. This condition means that expected highest of M draws exceeds the value M, when $M = h\left(\bar{\theta}\right)$.

Holding M fixed, the introduction of trade to the IEA increases the expected payoff of a signatory by

$$\frac{\sum_{i=k+1}^{M} \bar{\theta}_{(i,M)}}{M} - (M-k) \geq 0. \tag{5}$$

The inequality is strict if it is optimal to reduce required abatement (set $\bar{a} < 1$). The first term on the left side of inequality (5) is the expected cost savings and the second term is the certain reduction in abatement.

For fixed M, the introduction of trade has an ambiguous effect on global welfare. Trade allocates the level of abatement efficiently amongst signatories, but trade can reduce the equilibrium level of abatement, which is already too low to maximize global welfare. Trade reduces aggregate welfare (holding fixed the size of the IEA) if and only if

$$N(M-k) > \sum_{i=k+1}^{M} \bar{\theta}_{(i,M)}. \tag{6}$$

The term on the left side of this inequality is the global loss from the reduction in abatement, and the term on the right is the reduction in expected cost.

Trade also changes the incentive to participate in an IEA. In order to examine the stability of the IEA at the no-trade level of participation, $M = h\left(\bar{\theta}\right)$, denote k, $k^{(+1)}$ and $k^{(-1)}$ as the solutions to equation (3) for, respectively, $M = h\left(\bar{\theta}\right)$, $M = h\left(\bar{\theta}\right) + 1$ and $M = h\left(\bar{\theta}\right) - 1$. These three values of k are the equilibrium levels of total abatement if there are $M = h\left(\bar{\theta}\right)$ members, or one more member, or one few member. External stability requires that non-members do not want to join the agreement, and internal stability requires that members do not want to leave the agreement. Evaluated at the no-trade level of participation, the external and internal stability conditions are, respectively,

$$k^{(+1)} - \frac{\sum_{i=1}^{k^{(+1)}} \bar{\theta}_{(i,M^*)}}{M} < k \quad \text{and} \quad k^{(-1)} < k - \frac{\sum_{i=1}^{k} \bar{\theta}_{(i,M)}}{M}. \tag{7}$$

For example, with trade the IEA with $M = h\left(\bar{\theta}\right)$ members is not internally stable if

$$\frac{\sum_{i=1}^{k} \bar{\theta}_{(i,M)}}{M} > \left(k - k^{(-1)}\right),$$

which is likely to hold if a signatory's expected abatement cost under the original IEA is high and if one member's defection causes a small decrease in total abatement.

Equilibrium conditions that involve order statistics are difficult to evaluate for general distributions, so to obtain concrete results I specialize to the uniform distribution:

$$\theta \sim U\left[\theta_L, \theta_L + \varepsilon\right].$$

In this case, the no-trade equilibrium size of the IEA is

$$M = h\left(\theta_L + \frac{\varepsilon}{2}\right). \tag{8}$$

The expected i'th order statistic is

$$\bar{\theta}_{(i,M)} = \theta_L + \frac{i\varepsilon}{M+1}. \qquad * \tag{9}$$

(The Appendix contains derivations of equations with a "$*$" to the right of them.)

In order to obtain a simpler form of equation (3), define $g(x)$ as the largest integer not greater than x; $g(\cdot)$ is the "floor function". (A mnemonic: "g" for "ground".) Clearly, the number of members in any non-trivial equilibrium IEA must be greater than θ_L, so hereafter assume $M > \theta_L$. Set the right side of equation (9) equal to M and solve for i to obtain

$$i^* = \frac{M - \theta_L}{\varepsilon}(M+1). \tag{10}$$

The integer part of i^*, $g(i^*)$ equals the solution to equation (3) in the situation where that solution is less than M. Thus, the optimal level of abatement in an IEA with trade, given the uniform distribution, is

$$k = \min\left\{g\left(\frac{M - \theta_L}{\varepsilon}(M+1)\right), M\right\}. \tag{11}$$

This equation is the basis for demonstrating two aspects of trade. First, in many circumstances, trade gives non-members an incentive to join an agreement, because their participation increases the average abatement of existing members (Remark 2). Second, if the cost dispersion is high, the introduction of trade reduces the average level of abatement for a given level of membership (Proposition 1).

Remark 2 *Trade is positive if and only if the value of k given by equation (11) satisfies $0 < k < M$ When this inequality holds, the necessary and sufficient condition for an additional member to increase average abatement of existing members is*

$$g\left(\frac{M+1-\theta_L}{\varepsilon}(M+2)\right) \geq g\left(\frac{M-\theta_L}{\varepsilon}(M+1)\right) \qquad * \tag{12}$$

A sufficient (but not necessary) condition for inequality (12) is that the argument of the function of the left side exceeds the argument of the function on the right side by at least 1, a condition that reduces to

$$2(M+1) - (\varepsilon + \theta_L) > 0. \tag{13}$$

The sufficient condition for an additional member to increase abatement by existing members, inequality (13), is easy to check and clearly holds for a large part of parameter space. The necessary and sufficient condition, inequality (12) is easy to check numerically. I have not found any parameters for which that condition fails (when $0 < k < M$), so the conclusion that trade creates an incentive to join the IEA, in order to increase abatement by existing members, appears robust.

Proposition 1 *(i) Holding M fixed, trade reduces the equilibrium amount of abatement if and only if $\bar{\theta}_{(M,M)} > M$. With uniformly distributed costs, this inequality is equivalent to*

$$C \equiv \theta_L + \frac{M\varepsilon}{M+1} - M > 0. \tag{14}$$

(ii) A sufficient (but not necessary) condition for $C > 0$ (so that trade reduces equilibrium abatement, holding M fixed) evaluated at the no-trade IEA $M = h\left(\theta_L + \frac{\varepsilon}{2}\right)$ is

$$\varepsilon > 1 - \theta_L + \sqrt{\theta_L^2 + 2\theta_L + 9}. \tag{15}$$

A necessary condition for inequality (15) is $\epsilon > 2\sqrt{3}$.

(The Appendix contains all proofs.)

In the absence of trade, there is a unique non-trivial IEA. I am unable to analytically show uniqueness with trade. However simulations do not uncover any examples of multiple equilibria, so uniqueness appears to be robust. Lemma 2 in the Appendix provides a formula for the smallest internally stable non-trivial IEA with trade, and simulations support the conjecture that this equilibrium is also externally stable and is unique. If the cost dispersion (measured by ε) is sufficiently large, trade reduces the size of the minimum stable IEA:

Proposition 2 *A sufficient condition for the minimum (non-trivial) internally stable IEA under trade to be smaller than the unique (non-trivial) stable IEA without trade is*

$$\varepsilon > 3 - \theta_L + \sqrt{\theta_L^2 - 2\theta_L + 9}. \tag{16}$$

Example 1 *A numerical example shows that trade reduces the equilibrium size of the IEA even when the cost dispersion is much smaller than the level given in Proposition 2, and also that trade can reduce global welfare. For $\theta_L = 5$, the right side of equation (16) equals 2.9. However, Lemma 2 in the Appendix implies that the smallest internally stable IEA is $M = 6$ for $\varepsilon \in (0, 7)$ and $M = 7$ for $\varepsilon \in (7, 16)$. In contrast, the unique non-trivial IEA without trade, using equation (1), is $h(5 + 0.5\varepsilon)$; this value increases from 6 to 13 as ε ranges from slightly above 0 to slightly below 16. For this example, the IEA with trade is smaller than the IEA without trade whenever $\varepsilon > 2$; the bound given in Proposition 2 is not tight. Now consider the situation where ε is slightly below 16, where trade reduces equilibrium membership from 13 to 7. In this neighborhood, the average amount of abatement under trade (using equation (11)) is $\frac{1}{7}$. In this case, trade lowers abatement from 13 to 1, leading to a loss of global benefits of $12N$. Expected costs fall by*

$$13\bar{\theta} - \bar{\theta}_{(1,7)} = 13\left(5 + \frac{\varepsilon}{2}\right) - \left(5 + \frac{\varepsilon}{7+1}\right) = 60.0 + 6.375\varepsilon \approx 162$$

(for $\varepsilon \approx 16$). Net global benefits fall whenever $N > \frac{162}{12} = 13.5$. Since there are 13 members in the IEA without trade, global welfare falls unless the IEA without trade consists of all potential members (i.e. $N = 13$).

In the absence of trade, a signatory's defection from the no-trade IEA causes the remaining members to cease abatement. This credible threat supports the internal stability of the no-trade IEA. The following Proposition shows that trade eliminates this credible threat and thereby undermines internal stability.

Proposition 3 *If one signatory defects from the no-trade IEA, and the post-defection IEA allows trade, the remaining signatories choose a positive level of abatement strictly less than capacity. That is, the post-defection IEA involves positive abatement and trade.*

The extreme non-linearity and non-monotonicity of the benefit function makes it difficult to obtain further analytic results. However, because the model contains only two parameters, θ_L and ε, it is easy to obtain a complete numerical characterization of the equilibria.

Numerical analysis of cap and trade If the cost dispersion is sufficiently small, trade obviously has no effect on equilibrium abatement, but if the dispersion is large enough, trade necessarily decreases the equilibrium level of abatement. Figure 2 illustrates the possibility that the effect of cost dispersion on the level of abatement is non-monotonic. The functions graphed in this figure are evaluated at the no-trade equilibrium level of M, given by equation (8).

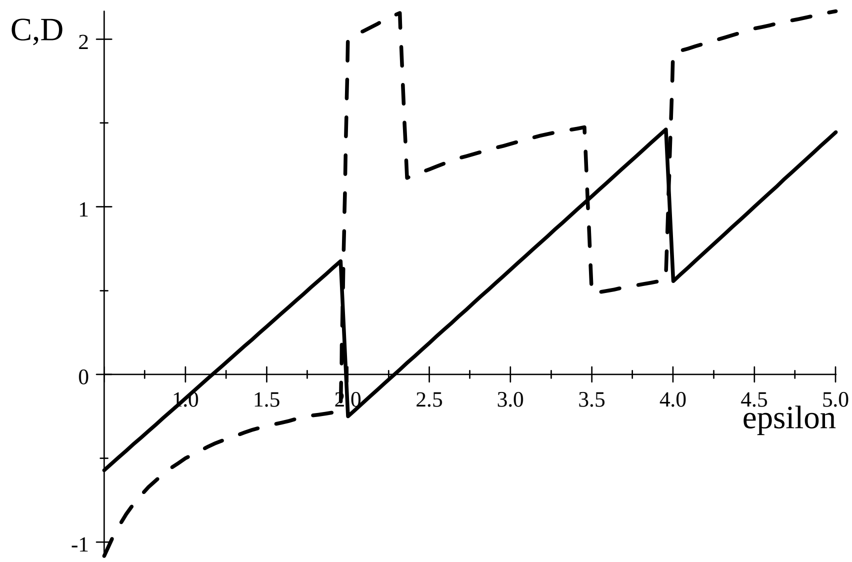

Figure 2: Graph of C (solid) and D (dashed) for $\theta_L = 5$. Trade reduces abatement in original no-trade IEA (holding M fixed) for $C > 0$, and trade reduces the size of the IEA for $D > 0$

The solid graph in this figure shows C (defined in equation (14)) as a function of ε for $\theta_L = 5$; holding membership fixed, trade reduces abatement for $C > 0$ and has no effect for $C \leq 0$. Trade has no effect on abatement for $\varepsilon < 1.2$ or for $\varepsilon \in (2, 2.25)$, but for all other values of ε, trade reduces equilibrium abatement.[3]

Extensive simulations show that the "original" no-trade IEA, $M = h\left(\bar{\theta}\right)$, is always externally stable when trade is allowed: no non-signatory wants to join once trade is allowed. The original IEA is not internally stable when the cost dispersion is high. Using the uniform distribution and the second part of inequality (7), the necessary and sufficient condition for the no-trade IEA to be internally unstable once trade is allowed is:

$$D \equiv \frac{k\theta_L + \frac{1}{2}k\left(k+1\right)\frac{\varepsilon}{M+1}}{M} - k + k^{(-1)} > 0. \qquad * \tag{17}$$

If this inequality is satisfied, then at least one signatory would want to leave the no-trade IEA once trade is allowed. The dashed graph in Figure 2 shows D as a function of ε for $\theta_L = 5$. For this example, the graph of D (unlike that of C) crosses the horizontal axis only once. There are four intervals worth noting.

1. For $C < 0 \wedge D < 0$ (where $\varepsilon < 1.2$), the original IEA without trade remains stable once trade is allowed; however, in this equilibrium there is no scope for trade because it is optimal to require all signatories to continue to abate at capacity.

[3]The graphs are shown as continuous for viewing ease. In fact, the graphs are discontinuous at the points where they become vertical.

2. For $C > 0 \wedge D < 0$ (where $\varepsilon \in (1.2, 2)$), the no-trade IEA remains stable once trade is allowed, but the optimal level of abatement falls.

3. For $C < 0 \wedge D > 0$ (where $\varepsilon \in (2, 2.25)$), if the no-trade IEA were to remain intact, it would be optimal to require all signatories to abate at capacity even if trade were allowed; however, the possibility of trade causes the IEA to shrink, and this change causes a reduction in per-signatory average abatement.

4. For $C > 0 \wedge D > 0$ (where $\varepsilon > 2.25$) the ability to trade causes the IEA to shrink relative to the no-trade level, leading to a fall in per-signatory average abatement. Even if the no-trade IEA could be held intact, the possibility of trade would reduce the average per-signatory abatement.

An appendix, available on request, provides additional numerical analysis. For sufficiently large θ_L, the graph of D crosses the 0 axis only once, at $\varepsilon = 2$. For these values of θ_L, the no-trade IEA remains stable once trade is allowed, if and only if $\varepsilon < 2$. However, for smaller values of θ_L, e.g. $\theta_L \in [1.2, 1.5]$ there are three values of ε at which $D = 0$. When $\theta_L \in [1.2, 1.5]$, for small values of ε the no-trade IEA is stable and for large values it is unstable, but as ε varies over intermediate ranges, the no-trade IEA switches between between being stable and non-stable, once trade is allowed. In this sense, stability is not monotonic in ε. This type of non-monotonicity illustrates the difficulty in obtaining a complete analytic characterization of the equilibrium.

3.2 The IEA with a safety valve

This section studies the effect of using a safety valve, i.e. a maximum price of permits, denoted $\bar{p}$. Because individual costs are not verifiable, implementing a price ceiling requires that countries are able to trade emissions permits. If the market price exceeds the price ceiling, signatories receive enough additional permits to drive the price to the ceiling. Alternatively, signatories are able to buy permits from a "carbon banker" at the price $\bar{p}$, and the banker returns the revenue from sales to signatories in a lump sum. The IEA also sets the per-member average abatement requirement, $\bar{a}$. As in the previous section, countries are price takers in the market for emissions permits. If the IEA sets the price ceiling before the participation decision, potential signatories take the ceiling as given. If the IEA conditions the price ceiling on the level of membership, a potential signatory understands that its participation decision can affect the price ceiling, thereby affecting the equilibrium level of abatement.

Figure 3 shows a particular realization of costs, and values of M, $\bar{p}$, and $\bar{a}$, in a situation where the price ceiling is binding. For this realization of costs and values of M and $\bar{a}$, the

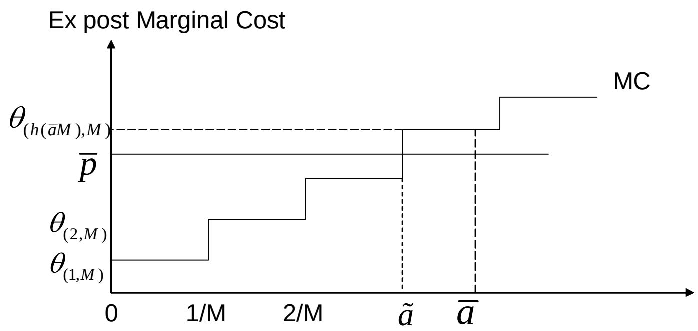

Figure 3: A cost realisation for which the price ceiling is binding

market price in the absence of the price ceiling would exceed $\bar{p}$. The price ceiling causes the carbon banker to release enough permits to drive the price to $\bar{p}$ or lower. There is indeterminacy in the price, because the marginal cost curve is a step function. Any price between $\bar{p}$ and the cost of the highest abater is an equilibrium price. However, this indeterminacy has no effect on the analysis. In the example shown in the figure, the resulting equilibrium average level of abatement is $\tilde{a}$, so the banker sells or distributes $M(\bar{a} - \tilde{a})$ additional permits.

For $\bar{a} = 1$, inspection of Figure 3 shows that the $k \leq M$ countries with the lowest costs abate at capacity, where[4]

$$k = \max_{0 \leq i \leq M} i \text{ subject to } \theta_{(i,M)} \leq \bar{p}. \tag{18}$$

The equilibrium k takes a particular value, say j, if and only if $\theta_{(j,M)} \leq \bar{p} < \theta_{(j+1,M)}$, which is equivalent to the event that among the M cost realizations, exactly j of them are lower than or equal to $\bar{p}$ and $M - j$ of them are higher than $\bar{p}$.

3.2.1 Choice of $\bar{p}$ before participation decision

Remark 1 notes that if the IEA chooses the abatement target before the participation decision, the unique equilibrium to the participation game is $M = 0$, regardless of whether the IEA allows trade amongst signatories. The analogous result holds with a price ceiling.

The key to the proof is showing that for a general value of $\bar{p}$ set prior to membership decisions, and with k given by equation (18), the expected value of per country average abatement, $a = \frac{k}{M}$, is independent of M. This independence means that by joining the IEA a country would have no effect on the expected abatement levels of other members. A country's decision to join an IEA would increase total expected abatement only by the amount of its own expected

[4]Define $\theta_{(0,M)} = 0$, so that if all costs exceed $\bar{p}$ no countries abate.

abatement. However, by virtue of the assumption $\theta_L > 1$, countries' want to abate only if by doing so they are able to increase the expected abatement of other countries. Therefore, when the IEA chooses $\bar{p}$ before countries make the participation decision, a potential signatory exerts no leverage over other signatories:

Proposition 4 *If the IEA chooses the price ceiling $\bar{p}$ before the participation stage, and sets $\bar{a} = 1$, the unique Nash equilibrium to the participation game is $M = 0$.*

3.2.2 Choice of $\bar{p}$ after participation decision

When the IEA chooses the price ceiling after membership, the choice $\bar{p} = M$ means that marginal abatement costs equal marginal benefits at every interior solution, i.e., when average abatement is less than capacity. The choice $\bar{a} = 1$ in this circumstance means that there are never unexploited beneficial abatement opportunities. Figure 3 shows a level of $\bar{a} < 1$ at which there would be unexploited beneficial abatement opportunities, if for example M happened to be greater than $\theta_{(h(\bar{a})M,M)}$.

Although the IEA has two instruments, $\bar{a}$ and $\bar{p}$, the first is redundant because it is optimal to set $\bar{a} = 1$ for any $M > \theta_L$. Thus, the IEA has a single instrument both with the cap and trade policy and with the price ceiling policy.

Remark 3 *The choice $\bar{a} = 1$ and $\bar{p} = M$ leads to the full-information ex post optimum. The "price policy" implemented by the safety valve is equivalent to a tax. For every realization, it is preferred to the "quantity policy" implemented by cap and trade without the price ceiling (Section 3.1). With the price policy, marginal cost equals marginal benefit in every realization where the capacity constraint is not binding; in contrast, with the quantity policy, marginal cost equals marginal benefit only in expectation.*

These results depend on the assumption that individual agents' marginal benefit of abatement are constant. For decreasing marginal benefits of abatement, the cap and trade with a price ceiling is an example of a hybrid policy, and it is not equivalent to a tax.[5]

For any distribution with bounded support, all price ceilings equal to or greater than the highest possible cost ($\bar{p} = \theta_L + \varepsilon$ under the uniform distribution) lead to the same outcome.

[5] The assumption that individual countries have constant marginal benefit of abatement means that the IEA's marginal benefit of abatement, conditional on membership, is a constant. This fact implies that with a price ceiling, the optimal cap equals the capacity level; consequently, the cap-and-trade with a price ceiling is equivalent to a tax. Moreover, the fact that marginal benefits are constant means that the optimal tax leads to higher expected welfare than the optimal quantity policy, for reasons well understood from the "prices versus quantities" literature.

Therefore, for the uniform distribution there is no loss in generality in assuming that $\bar{p} \leq \theta_L + \varepsilon$. With this convention, the ex post *optimal* price ceiling is

$$\bar{p} = \min\{\theta_L + \varepsilon, M\}. \tag{19}$$

Equation 19 implies that the optimal price ceiling – and thus, expected abatement – weakly increases with membership. Therefore, a potential signatory has an incentive to join the agreement, in order to increase the abatement of existing members. However, the use of an ex post optimal price ceiling decreases the severity of the consequences of leaving (or not joining) an IEA. On balance, the price ceiling, like cap and trade, decreases equilibrium membership.

With M members, the expected aggregate abatement equals

$$M \times \min\left\{\frac{M - \theta_L}{\varepsilon}, 1\right\}. \quad * \tag{20}$$

The following lemma simplifies expression (20) by providing a sufficient condition to ensure that $M \leq \theta_L + \varepsilon$ in equilibrium.

Lemma 1 *A sufficient condition for $M \leq \theta_L + \varepsilon$ in equilibrium is $\varepsilon \geq 2$.*

The equilibrium value of M may depend on whether the IEA allows only trade or uses trade together with a price ceiling. As noted above, the trade-only policy corresponds to a cap and trade regime, and the price ceiling policy corresponds to a tax regime. The expected abatement might be higher or lower under the price ceiling (equivalently, the tax regime), relative to the cap and trade regime. In order to establish this claim in a simple manner, I restrict attention to parameter space $\varepsilon \geq 2$ and $M \leq \theta_L + \varepsilon$. Denote the difference between total expected abatement under the price ceiling and abatement under the quantity target, conditional on M, as δ. The following holds:

Proposition 5 *(i) For $\varepsilon \geq 2$ and given $M \leq \theta_L + \varepsilon$, the difference between total expected abatement under the price ceiling and abatement under the quantity target satisfies the bounds*

$$-\frac{M - \theta_L}{\varepsilon} \leq \delta < 1 - \frac{M - \theta_L}{\varepsilon}. \tag{21}$$

(ii) For $\varepsilon \geq 2$ and given $M \leq \theta_L + \varepsilon$, δ can be positive or negative.

In addition:

Proposition 6 *For $\varepsilon > 2$, when the no-trade IEA ($M = h\left(\bar{\theta}\right)$) begins to use the optimal price ceiling, expected abatement falls.*

Propositions 1 and 6 show that trade, with or without the price ceiling, both reduce (expected) abatement when the cost dispersion is large enough.

In order to determine how the price ceiling affects the equilibrium size of the IEA, I consider the condition for internal stability. When $\varepsilon > 2$, so that (by Lemma 1) $M \leq \theta_L + \varepsilon$, the condition for internal stability of an IEA with M members, when the IEA uses the ex post optimal price ceiling, is

$$\frac{1}{2}\frac{M-\theta_L}{\varepsilon}\left((M+\theta_L)-\left(\frac{M-\theta_L}{\varepsilon}\right)^M(\theta_L+\varepsilon-M)\right)-\frac{2M-1-\theta_L}{\varepsilon}<0. \quad * \quad (22)$$

Numerical analysis of the price ceiling Using equation (22), I obtain a three dimensional graph showing the net gain of defection from the no-trade IEA ($M = h\left(\bar{\theta}\right)$), once that IEA begins to use the price ceiling, as a function of ε and θ_L. For $\varepsilon \geq 2.01$ and $\theta_L \geq 1.01$, there is a very small region of parameter space, with small values of ε and θ, where the net gain to defection is negative. In that region, the no-trade IEA is stable under the price ceiling, even though the price ceiling does reduce abatement. However, for "most" of parameter space, the net gain from defection is positive, so the no-trade IEA is not internally stable when the IEA begins to use a price ceiling.

Comparison with Example 1, where $\theta_L = 5$, provides a simple illustration of the effect of using the safety valve. For $\varepsilon < 2$ the unique non-trivial IEA consists of six members for all three policy regimes: no trade, cap and trade, or a price ceiling. For $\varepsilon < 2$, no country wants to defect from a six-member IEA because doing so causes abatement to fall to 0. However, when the IEA uses a price ceiling, *the only stable IEA consists of six members, for all values* $\varepsilon > 0$. The IEA without trade contains $h\left(\bar{\theta}+\frac{\varepsilon}{2}\right)$ and Example 1 notes that under cap and trade without the price ceiling, the IEA has seven members for $\varepsilon \in (7, 16)$; that IEA has more than seven members for $\varepsilon > 16$.

In summary, the introduction of trade, with or without the price ceiling, reduces the equilibrium IEA size (relative to the no-trade regime) whenever there is sufficient cost dispersion ($\varepsilon > 2$). When the cost dispersion is large (here, $\varepsilon > 7$), the price ceiling decreases the number of IEA members, relative to the trade-only regime. For small cost dispersion, use of the price ceiling does not change the trade-only equilibrium IEA size. For intermediate levels of cost dispersion ($2 < \varepsilon < 7$), where the equilibrium IEA has six members under trade-only or trade-with-price-ceiling, expected abatement might be larger under either regime.

4 Other mechanisms

Society has two objectives: to induce countries to join the IEA, and to induce signatories whose marginal cost is below social marginal benefit to abate. For almost any model, it would be possible to design a mechanism that achieves these objectives. However, those complicated mechanisms may have limited policy relevance. This paper focuses on simple institutions, or mechanisms. Recall that, in the context of this model, a country's incentive to join depends on the effect that its participation has on the actions of other members.

In the absence of trade, there is a critical membership level ($M = h(\bar{\theta}) - 1$) at which a country has a substantial incentive to join, but for all other membership levels it has no incentive to join. The resulting equilibrium membership is low. Conditional on membership $M = h(\bar{\theta})$, countries with high abatement costs may be required to abate. Therefore, the no-trade setting does a poor job in achieving both of society's objectives.

Cap and trade, where the IEA conditions the cap on membership level, does a reasonably good job of determining which members actually abate. However, relative to no-trade, cap and trade leads to even less powerful participation incentives. A potential member knows that its decision not to join leads to slightly lower, rather than a much lower, level of members' abatement. On balance, the cap and trade regime is likely to lower expected social welfare when the cost dispersion is high – exactly the circumstance when we expect trade to be most useful.

Relative to cap and trade, the safety valve that is chosen conditional on membership level does a better job of selecting which members actually abate; but it leads to an even lower-powered incentive to join the IEA. Equilibrium welfare with the safety valve tends to be even lower than under cap and trade.

The safety valve chosen prior to membership – like the no-trade and the cap and trade abatement levels chosen prior to membership – eliminates the incentive to join the IEA. These three ex ante mechanisms induce a unique equilibrium with no members, and consequently no abatement.

At the risk of potentially sacrificing efficiency conditional on membership, we can chose a safety valve mechanism that promotes membership. Figure 4 shows three graphs of price ceilings as a function of M, ignoring here that M is an integer. (The points of intersection of the three graphs are not relevant to the discussion here.) The price ceiling chosen before membership is independent of M and eliminates all incentive to join the agreement. The ex post optimal price ceiling, equation (19), graphed as the dashed 45^o line, has some "power". It increases with M, so potential members know that by joining they increase the price ceiling and thereby tend to increase other signatories' abatement. However, the incentive is low-powered.

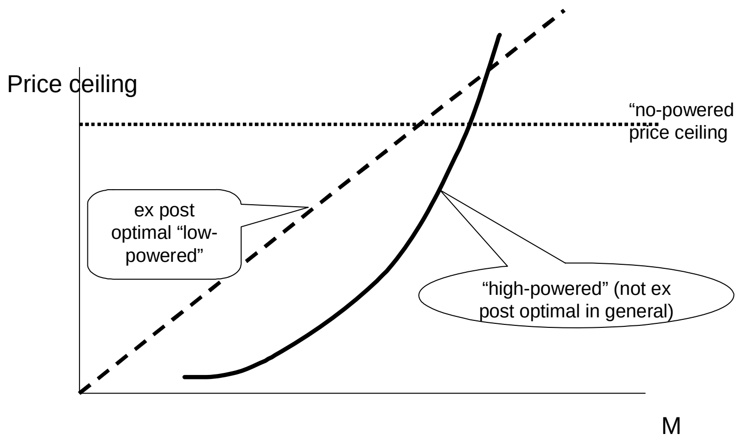

Figure 4: The "no-powered" constant price ceiling, chosen before membership (dotted). The "low-powered" ex post optimal price ceiling, chosen conditional on membership (dashed). The "high powered" price ceiling, a function chosen before membership (solid).

Average abatement is still rather high even when membership is low, and increased membership leads to a rather small increase in the price ceiling, and therefore in abatement, in the region near the equilibrium – the region that matters.

The solid convex curve in Figure 4 shows a "high-powered" safety valve. The key features of this mechanism are that it lies below the graph of the ex post optimal price ceiling for low membership levels, but at some membership level it begins to increase rapidly. For low membership levels, the price ceiling is strictly below the ex post optimal level. For these levels, members abate less than the level that maximizes their group welfare. However, where the price ceiling begins to increase rapidly, potential members have a significant incentive to join the IEA, because by doing so they significantly increase the average abatement of other members. For example, suppose that the number of potential members, N, satisfies $N < \theta_L + \varepsilon$; that is, the cost dispersion is great enough that for some realizations the marginal cost of abatement exceeds the marginal benefit even when all countries join the IEA. In that case, there is a first best safety valve mechanism, $\bar{p}(M)$, that sets $\bar{p}(N) = N$ and $\bar{p}(M) << M$ for $M < N$. All countries want to join, because failure to do so would lead to a low price ceiling and low abatement levels by members; in equilibrium only countries with costs lower than benefits actually abate.

This high-powered safety valve is similar to the escape clause studied in Karp and Zhao (2010). In the simplest version of the escape clause, countries negotiate a required level of abatement and a "fine", F *prior* to the participation game. Signatories to the IEA are in compliance if they either achieve the required level of abatement or pay the fine. All fine revenues are distributed equally among signatories that are in compliance. If there are M

signatories, the fine net of the reimbursement (the "effective fine") is $F\left(1-\frac{1}{M}\right)$, which is increasing in M. Thus, an additional member increases the effective fine for all other members, making it less attractive for them to exercise the escape clause, and more likely that they will abate. The escape clause therefore provides a high-powered incentive to join the IEA.

The high-powered safety valve proposed here also has an interesting relation to earlier proposals that the IEA (somehow) commit to a "consensus treaty" or to behaving "modestly". With such an agreement, the IEA internalizes only a fraction of the group benefits of abatement, e.g. by choosing an emissions tax equal to a fraction (less than 1) of marginal group damages. Barrett (2002) and Finus and Maus (2008) show that this structure can increase equilibrium membership and global welfare, relative to the situation where the IEA behaves optimally conditional on membership. The high-powered price ceiling that I propose requires suboptimal abatement, but only at non-equilibrium levels of participation.

5 Conclusion

Countries have least four types of reasons to join an IEA: to influence the actions of other members; to influence the participation decisions of other potential signatories; to constrain their own future actions (a commitment device) and to influence the actions of non-signatories. By virtue of its assumptions, the model here incorporates only the first type of reason. The implications of the model have to be interpreted with this limitation in mind.

The first major conclusion is that a cap and trade policy with or without a safety valve reduces at least the first of the four reasons that a country has for joining the IEA. These policies may therefore lower rather than increase welfare. This conclusion is consistent with earlier results from the IEA literature, regarding the equilibrium effect of a decrease in abatement costs; but it does not follow from those results. Both trade policies create incentives going in the right direction. For example, under the price ceiling, by joining the IEA, the new member increases the equilibrium price ceiling and thereby increases existing members' abatement. However, that positive effect is not strong enough to offset the negative effect: under an ex post optimal price ceiling or the cap and trade, a member knows that its defection would cause a reduction in remaining members' abatement, but not a collapse of the IEA.

The second major conclusion is that a "high-powered" safety valve can promote membership. This policy chooses the relation between membership and the price ceiling at the negotiation stage, prior to the participation decisions. It requires that the price ceiling be lower than the ex post optimal level for low membership levels, but that it rise rapidly for high membership levels.

A Appendix: Derivations

Derivation of equation (9) If θ is *uniformly* distributed on $[\theta_L, \theta_H]$, the kth order statistics $\theta_{(k,M)}$ with a membership of M are distributed with the following density function:

$$\frac{M!}{(k-1)!(M-k)!}\left(\frac{\theta_{(k,M)}-\theta_L}{\varepsilon}\right)^{k-1}\left(1-\frac{\theta_{(k,M)}-\theta_L}{\varepsilon}\right)^{M-k}\frac{1}{\varepsilon}. \tag{23}$$

That is, $\frac{\theta_{(k,M)}-\theta_L}{\varepsilon}$ has a Beta distribution $B(\theta_{(k,M)}; k, M+1-k)$, with mean

$$E\left(\frac{\theta_{(k,M)}-\theta_L}{\varepsilon}\right)=\frac{k}{M+1}. \tag{24}$$

Equation (24) implies equation (9).

Derivation of inequality (12): The number of members M and the value of k given by equation (11) are both integers, and $0 < k < M$ by assumption. Sufficiency: For any two integers $M \geq 2$ and $0 < k < M$ we have

$$\frac{k}{M} < \min\left\{\frac{k+1}{M+1}, 1\right\}. \tag{25}$$

Inequality (12) implies that right side of inequality (25) is less than or equal to the average level of abatement when there are $M+1$ members. Necessity: If inequality (12) does not hold, then aggregate abatement remains constant or possibly falls with an additional member, so the average abatement of the original members must fall.

Derivation of inequality (17): The uniform distribution and the second part of inequality (7) imply that the original no-trade IEA is internally unstable iff

$$\begin{array}{c} D \equiv \frac{\sum_{i=1}^{k}\bar{\theta}_{(i,M)}}{M} - k + \hat{k} = \\ \frac{\sum_{i=1}^{k}\left(\theta_L + \frac{i\varepsilon}{M+1}\right)}{M} - k + \hat{k} = \frac{k\theta_L + \frac{\varepsilon}{M+1}\sum_{i=1}^{k} i}{M} - k + \hat{k} \\ = \frac{k\theta_L + \frac{1}{2}k(k+1)\frac{\varepsilon}{M+1}}{M} - k + \hat{k} > 0. \end{array}$$

Derivation of equation (20): This equation follows from equations (19) and (37), below.

Derivation of equation (22): Using equation (37), below, defection by one signatory decreases expected abatement per signatory and also the number of signatories, so the defection necessarily decreases expected abatement. This decrease in expected abatement is

$$\Lambda(M,\varepsilon) \equiv M \times \min\left\{1, \frac{M-\theta_L}{\varepsilon}\right\} - (M-1) \times \min\left\{1, \frac{M-1-\theta_L}{\varepsilon}\right\} > 0.$$

The defecting signatory also saves the expected cost $\Omega(M, \varepsilon, \bar{p})$, a function defined in equation (32) below, with $\bar{p}$ determined by equation (19). The net benefit of defection is therefore

$$F(M,\varepsilon) \equiv \Omega(M,\varepsilon,\bar{p}) - \Lambda(M,\varepsilon). \tag{26}$$

When $M \leq \theta_L + \varepsilon$, which by Lemma 1 is guaranteed for $\varepsilon \geq 2$, the reduction in abatement due to a one-country defection simplifies to

$$\Lambda(M,\varepsilon) = M\left(\frac{M-\theta_L}{\varepsilon}\right) - (M-1)\left(\frac{M-1-\theta_L}{\varepsilon}\right) = \frac{2M-1-\theta_L}{\varepsilon}.$$

Substituting this expression and the definition of $\Omega(M,\varepsilon,\bar{p})$ into equation (26) yields equation (22).

B Appendix: Proofs

Proof. (**Proposition 1**) Part (i) follows immediately from the comments above the proposition, so it requires no further proof. To prove part (ii) I use part (i) and the definition of C in equation (14), to conclude that trade reduces abatement at the no-trade IEA iff

$$\theta_L + \frac{M\varepsilon}{M+1} - M = \frac{1}{M+1}\left(-M^2 + (\varepsilon + \theta_L - 1)M + \theta_L\right) > 0.$$

Trade reduces abatement iff

$$\left(-M^2 + (\varepsilon + \theta_L - 1)M + \theta_L\right) > 0.$$

The quadratic on the left side is concave in M and positive at $M = 0$. The positive root of the quadratic is

$$\frac{1}{2}(\varepsilon + \theta_L - 1) + \frac{1}{2}\sqrt{\varepsilon^2 + 2\varepsilon\theta_L - 2\varepsilon + \theta_L^2 + 2\theta_L + 1},$$

so trade reduces abatement, evaluated at the no-trade IEA iff

$$M = h\left(\theta_L + \frac{\varepsilon}{2}\right) < \frac{1}{2}(\varepsilon + \theta_L - 1) + \frac{1}{2}\sqrt{\varepsilon^2 + 2\varepsilon\theta_L - 2\varepsilon + \theta_L^2 + 2\theta_L + 1}. \quad (27)$$

A sufficient (but not necessary) condition for this inequality is

$$\left(\theta_L + \frac{\varepsilon}{2}\right) + 1 < \frac{1}{2}(\varepsilon + \theta_L - 1) + \frac{1}{2}\sqrt{\varepsilon^2 + 2\varepsilon\theta_L - 2\varepsilon + \theta_L^2 + 2\theta_L + 1},$$

which is equivalent to

$$\begin{gathered}\left(\theta_L + \tfrac{\varepsilon}{2}\right) + 1 - \tfrac{1}{2}(\varepsilon + \theta_L - 1) - \tfrac{1}{2}\sqrt{\varepsilon^2 + 2\varepsilon\theta_L - 2\varepsilon + \theta_L^2 + 2\theta_L + 1} \\ = \tfrac{1}{2}\theta_L + \tfrac{3}{2} - \tfrac{1}{2}\sqrt{\varepsilon^2 + 2\varepsilon\theta_L - 2\varepsilon + \theta_L^2 + 2\theta_L + 1} < 0,\end{gathered}$$

or

$$\frac{1}{2}\theta_L + \frac{3}{2} < \frac{1}{2}\sqrt{\varepsilon^2 + 2\varepsilon\theta_L - 2\varepsilon + \theta_L^2 + 2\theta_L + 1}.$$

Now square both sides of this inequality to obtain the sufficient condition

$$\left(\frac{1}{2}\theta_L + \frac{3}{2}\right)^2 < \frac{1}{4}\left(\varepsilon^2 + 2\varepsilon\theta_L - 2\varepsilon + \theta_L^2 + 2\theta_L + 1\right).$$

Simplifying this inequality, we have

$$\frac{1}{2}(1-\theta_L)\varepsilon - \frac{1}{4}\varepsilon^2 + (\theta_L + 2) < 0.$$

The quadratic on the left side is concave in ε and positive at $\varepsilon = 0$, so the inequality holds if and only if ε is greater than the positive root, i.e. iff

$$\varepsilon > 1 - \theta_L + \sqrt{\theta_L^2 + 2\theta_L + 9}.$$

The right side of this inequality is increasing in θ_L and therefore obtains its greatest lower bound at the lower bound $\theta_L = 1$, where the right side of the inequality equals $2\sqrt{3}$. ■

In order to determine how trade affects the equilibrium size of the IEA, consider the circumstance where a no-trade IEA has already formed, and the IEA begins to allow trade. The stability of the resulting IEA depends on the consequence of defection by one signatory. If one signatory defects from the no-trade IEA, the level of abatement is $k^{(-1)}$, with

$$k^{(-1)} = \min\left\{ g\left(\frac{M-1-\theta_L}{\varepsilon}M\right), M-1 \right\}. \tag{28}$$

The following lemma gives the minimum internally stable IEA under trade:

Lemma 2 *(i) The necessary and sufficient condition for abatement to be positive when an IEA with M members allows trade is*

$$M > \frac{1}{2}\left(\theta_L - 1 + \sqrt{\theta_L^2 + 2\theta_L + 4\varepsilon + 1}\right). \tag{29}$$

(ii) An IEA with

$$M = h\left(\frac{1}{2}\left(\theta_L - 1 + \sqrt{\theta_L^2 + 2\theta_L + 4\varepsilon + 1}\right)\right) \tag{30}$$

members is the smallest internally stable IEA under trade.

Proof. (Lemma 2) (part (i)) Using (11), a necessary and sufficient condition for positive abatement is $k \geq 1$, or

$$g\left(\frac{M-\theta_L}{\varepsilon}(M+1)\right) \geq 1.$$

This inequality requires that the argument of $g(\cdot)$ is greater than or equal to 1, i.e.

$$(M-\theta_L)(M+1) - \varepsilon > 0$$

The quadratic in M is convex and negative at $M = 0$, so the inequality is satisfied if and only if M is greater than the positive root,

$$\frac{1}{2}\left(\theta_L - 1 + \sqrt{\theta_L^2 + 2\theta_L + 4\varepsilon + 1}\right).$$

To establish part (ii) note that an IEA equal to the size given in equation (30) abates at a positive level, giving signatories positive net benefit. A smaller IEA would choose not to abate, leaving a defector from the original IEA with 0 net benefits. Therefore, a signatory to the original IEA does not want to defect. ■

Proof. (Proposition 2) Comparison of the equations (1) and (30) shows that the minimum internally stable IEA with trade is less than the non-trivial stable IEA without trade if and only if

$$h\left(\frac{1}{2}\left(\theta_L - 1 + \sqrt{\theta_L^2 + 2\theta_L + 4\varepsilon + 1}\right)\right) < h\left(\theta_L + \frac{1}{2}\varepsilon\right).$$

A sufficient condition for this inequality is

$$\frac{1}{2}\left(\theta_L - 1 + \sqrt{\theta_L^2 + 2\theta_L + 4\varepsilon + 1}\right) < \theta_L + \frac{1}{2}\varepsilon - 1.$$

This inequality is equivalent to

$$\sqrt{\theta_L^2 + 2\theta_L + 4\varepsilon + 1} < \theta_L + \varepsilon - 1$$

Squaring both sides and rearranging, the inequality is equivalent to

$$(6 - 2\theta_L)\,\varepsilon - \varepsilon^2 + 4\theta_L < 0.$$

The quadratic on the left side is concave in ε and positive at $\varepsilon = 0$, so the inequality holds if and only if ε exceeds the positive root, which equals

$$3 - \theta_L + \sqrt{\theta_L^2 - 2\theta_L + 9}.$$

■

Proof. (Proposition 3) I first prove that abatement remains positive in the post-defection IEA, and then show that the post-defection IEA abates at less than capacity, so trade is positive.

If one country defects from the no-trade IEA and trade is allowed, the remaining number of countries is

$$M^* \equiv h\left(\theta_L + \frac{\varepsilon}{2}\right) - 1 \leq \theta_L + \frac{\varepsilon}{2} + 1 - 1 = \theta_L + \frac{\varepsilon}{2} \tag{31}$$

A necessary and sufficient condition for the IEA with M^* to continue abating is for M^* to satisfy inequality (29). A sufficient condition for that to hold, using equation (31) is

$$\theta_L + \frac{\varepsilon}{2} > \frac{1}{2}\left(\theta_L - 1 + \sqrt{\theta_L^2 + 2\theta_L + 4\varepsilon + 1}\right),$$

which with some rearrangement is equivalent to

$$(\theta_L + \varepsilon + 1)^2 > \sqrt{\theta_L^2 + 2\theta_L + 4\varepsilon + 1}$$
$$\Longleftrightarrow$$
$$\varepsilon\,(2\theta_L + \varepsilon - 2) > 0.$$

The last inequality always holds because $\theta_L > 1$.

In order to show that expected trade is positive in the post-defection IEA, I need to establish that inequality (14) is satisfied at $M = h\left(\theta + \frac{\varepsilon}{2}\right) - 1$. Repeating the argument that leads to inequality (27), the necessary and sufficient condition becomes

$$M = h\left(\theta_L + \frac{\varepsilon}{2}\right) - 1 < \frac{1}{2}\left(\varepsilon + \theta_L - 1\right) + \frac{1}{2}\sqrt{\varepsilon^2 + 2\varepsilon\theta_L - 2\varepsilon + \theta_L^2 + 2\theta_L + 1}.$$

The sufficient condition is

$$\left(\theta_L + \frac{\varepsilon}{2}\right) - 1 - \frac{1}{2}\left(\varepsilon + \theta_L - 1\right) = \frac{1}{2}\left(\theta_L - 1\right) < \frac{1}{2}\sqrt{\varepsilon^2 + 2\varepsilon\theta_L - 2\varepsilon + \theta_L^2 + 2\theta_L + 1}.$$

Squaring both sides of the last inequality and simplifying, we have the sufficient condition

$$\left(\frac{1}{2} - \frac{1}{2}\theta_L\right)\varepsilon - \frac{1}{4}\varepsilon^2 - \theta_L < 0,$$

which is always true because $\theta_L > 1$. ■

Proposition 4 relies on the following lemma:

Lemma 3 *When costs are uniform and the signatory uses a price ceiling $\bar{p} \leq \theta_L + \varepsilon$, expected costs for each signatory are*

$$\Omega\left(M, \bar{p}\right) = \frac{1}{2}\frac{\bar{p} - \theta_L}{\varepsilon}\left[\left(\bar{p} + \theta_L\right) - \left(\frac{\bar{p} - \theta_L}{\varepsilon}\right)^M \left(\theta_L + \varepsilon - \bar{p}\right)\right]. \tag{32}$$

Proof. (Lemma 3). Given the value of k (the number of countries that abate at capacity), determined by equation (18), the carbon banker releases $M - k$ additional permits, either selling them or distributing them freely. I present the proof for the case where the carbon banker distributes the permits freely. Under the alternative, the banker sells (at price $\bar{p}$) the additional permits, and then distributes the revenue, $\bar{p}\left(M - k\right)$, to members in a lump sum. Since the additional aggregate cost to members equals the additional aggregate revenue, and since each firm has the same probability of a particular rank in costs, the expected net cost per member is the same regardless of whether the banker sells or freely distributes the permits.

Under the assumption of independently and identically uniformly distributed costs, for arbitrary $\bar{p} \leq \theta_L + \varepsilon$, the probability that j countries abate to $\bar{a} = 1$ (i.e., the probability that $k = j$) is

$$\Pr\left\{k = j\right\} = \binom{M}{k}\left(\frac{\bar{p} - \theta_L}{\varepsilon}\right)^j \left(\frac{\theta_L + \varepsilon - \bar{p}}{\varepsilon}\right)^{M-j}. \tag{33}$$

To calculate expected costs, conditional on M and $\bar{p}$, I need to consider two situations: where the price ceiling is not binding and where it is binding. The probability that the price ceiling is not binding equals the probability that $k = M$; using equation (33), this probability equals

$$\alpha \equiv \left(\frac{\bar{p} - \theta_L}{\varepsilon}\right)^M . \tag{34}$$

Conditional on the price ceiling not being binding, each country has a cost realization less than $\bar{p}$. In this case, the conditional distribution of a country's cost is uniform over $[\theta_L, \bar{p}]$ and the expected cost conditional on this event is $\frac{\bar{p}+\theta_L}{2}$.

I now determine the expected costs conditional on the price ceiling being binding. In this case, the required amount of abatement per member is $\frac{k}{M}$, where the distribution of k is given by equation (33). If a country gets a realization $\theta < \bar{p}$ it spends θ on abatement and it sells $1 - \frac{k}{M}$ abatement credits. Its net cost in this case is $\theta - \left(1 - \frac{k}{M}\right)\bar{p}$. In writing this expression I assume that when the ceiling is binding, the market price equals $\bar{p}$. Inspection of Figure 3 shows that there is a range of equilibrium prices in this case, $\left[\theta_{(k,M)}, \bar{p}\right]$. The selection of $\bar{p}$ rather than a lower value as the equilibrium price is merely a convention. It transfers surplus from buyers to sellers. However, since every country has the same probability of a particular cost ranking (any two countries have the same probability of being a buyer or seller), the transfer has no effect on a country's expected costs.

If a country has a realization $\theta > \bar{p}$ its cost is $\frac{k}{M}\bar{p}$. Conditional on k, a signatory's expected abatement cost is

$$\begin{aligned} &\tfrac{1}{\varepsilon}\left(\int_{\theta_L}^{\bar{p}}\left(\theta - \left(1 - \tfrac{k}{M}\right)\bar{p}\right)d\theta + \int_{\bar{p}}^{\theta_L+\varepsilon}\tfrac{k}{M}\bar{p}d\theta\right) \\ &= \tfrac{k}{M}\bar{p} - \tfrac{(\bar{p}-\theta_L)^2}{2\varepsilon}. \end{aligned} \tag{35}$$

I now find the expectation of $\frac{k}{M}$, conditional on $k < M$. This expectation is

$$\begin{aligned} E(\left(\tfrac{k}{M} \mid k < M\right) &= \tfrac{1}{1-\alpha}\sum_{k=0}^{k=M-1}\binom{M}{k}\left(\tfrac{\bar{p}-\theta_L}{\varepsilon}\right)^k\left(\tfrac{\theta_L+\varepsilon-\bar{p}}{\varepsilon}\right)^{M-k}\tfrac{k}{M} \\ &= \tfrac{1}{1-\alpha}\left(\sum_{k=0}^{k=M}\binom{M}{k}\left(\tfrac{\bar{p}-\theta_L}{\varepsilon}\right)^k\left(\tfrac{\theta_L+\varepsilon-\bar{p}}{\varepsilon}\right)^{M-k}\tfrac{k}{M} - \binom{M}{M}\left(\tfrac{\bar{p}-\theta_L}{\varepsilon}\right)^M\left(\tfrac{\theta_L+\varepsilon-\bar{p}}{\varepsilon}\right)^{M-M}\tfrac{M}{M}\right) \\ &= \tfrac{1}{1-\alpha}\left(\tfrac{\bar{p}-\theta_L}{\varepsilon} - \alpha\right), \end{aligned} \tag{36}$$

where the last line uses equation (37) and the definition of α.

Using the results above, the unconditional cost to a member of an IEA with M members is

$$\alpha\frac{\bar{p}+\theta_L}{2}+(1-\alpha)\left[\left(\frac{1}{1-\alpha}\left(\frac{\bar{p}-\theta_L}{\varepsilon}-\alpha\right)\bar{p}\right)-\frac{1}{2}\frac{(\bar{p}-\theta_L)^2}{\varepsilon}\right]$$

$$=\alpha\left(\frac{\bar{p}+\theta_L}{2}-\bar{p}\right)+\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)\bar{p}-(1-\alpha)\frac{1}{2}\frac{(\bar{p}-\theta_L)^2}{\varepsilon}$$

$$=\alpha\left(\frac{\theta_L-\bar{p}}{2}+\frac{1}{2}\frac{(\bar{p}-\theta_L)^2}{\varepsilon}\right)+\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)\bar{p}-\frac{1}{2}\frac{(\bar{p}-\theta_L)^2}{\varepsilon}$$

$$-\frac{1}{2}\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)^M\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)(\theta_L+\varepsilon-\bar{p})+\frac{1}{2}\frac{\bar{p}-\theta_L}{\varepsilon}(\bar{p}+\theta_L)$$

$$\frac{1}{2}\frac{\bar{p}-\theta_L}{\varepsilon}\left[(\bar{p}+\theta_L)-\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)^M(\theta_L+\varepsilon-\bar{p})\right]$$

Using the definition of α in equation (34) completes the proof. ■

Proof. (Proposition 4) Using the density in equation (33), the expected value of per-country average abatement $\frac{k}{M}$ is[6]

$$E\left(\frac{k}{M}\right)=\sum_{k=0}^{k=M}\binom{M}{k}\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)^k\left(\frac{\theta_L+\varepsilon-\bar{p}}{\varepsilon}\right)^{M-k}\frac{k}{M}=\frac{\bar{p}-\theta_L}{\varepsilon}\leq 1. \qquad (37)$$

Using equations (37) and (32), the signatory's expected payoff is

$$\pi^S(M;\bar{p})=\frac{\bar{p}-\theta_L}{\varepsilon}M-\frac{1}{2}\frac{\bar{p}-\theta_L}{\varepsilon}\left[(\bar{p}+\theta_L)-\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)^M(\theta_L+\varepsilon-\bar{p})\right]$$

$$\frac{\bar{p}-\theta_L}{\varepsilon}\left(M-\frac{1}{2}\left[(\bar{p}+\theta_L)-\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)^M(\theta_L+\varepsilon-\bar{p})\right]\right)$$

.

The expected payoff of being a non-signatory when $M-1$ countries join, is

$$\pi^N(M-1;\bar{p})=\frac{\bar{p}-\theta_L}{\varepsilon}(M-1).$$

The advantage of joining an IEA that currently has $M-1$ members equals the difference between the welfare of a signatory in an IEA with M members and of a non-signatory when the IEA has $M-1$ members:

$$\pi^S(M;\bar{p})-\pi^N(M-1;\bar{p})=$$

$$\frac{\bar{p}-\theta_L}{\theta_H-\theta_L}\left(M-\frac{1}{2}\left[(\bar{p}+\theta_L)-\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)^M(\theta_L+\varepsilon-\bar{p})\right]-(M-1)\right)=$$

$$\frac{\bar{p}-\theta_L}{\varepsilon}\times G(\bar{p},M;\varepsilon,\theta_L)$$

where

$$G(\bar{p},M;\varepsilon,\theta_L)\equiv\left(1-\frac{1}{2}\left[(\bar{p}+\theta_L)-\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)^M(\theta_L+\varepsilon-\bar{p})\right]\right)$$

[6] As noted above, when $M>\theta_L+\varepsilon$, the choice of either $\theta_L+\varepsilon$ or M as the price ceiling leads to the same outcome; but the validity of equation (37) requires that $\bar{p}\leq\theta_L+\varepsilon$.

The function G is strictly decreasing in M. Consequently, there is at most one value of M that satisfies $G = 0$. Denote this root as M^*. Setting $G = 0$ and solving for M gives

$$M^* = \frac{\ln\left(\frac{(\bar{p}+\theta_L)-2}{(\theta_L+\varepsilon-\bar{p})}\right)}{\ln\left(\frac{\bar{p}-\theta_L}{\varepsilon}\right)} \tag{38}$$

The only candidate for a positive Nash equilibrium to the participation game is the largest integer not greater than M^*. Denote this integer as M^{**}. (At any M equal to an integer strictly greater than M^*, M is not internally stable, because an IEA member does better by becoming a non-signatory. At any integer M strictly less than M^{**}, a non-signatory gains more than $\frac{\bar{p}-\theta_L}{\theta_H-\theta_L} \times g\left(\bar{p}, M^{**}; \theta_H, \theta_L\right) > 0$ by joining, so any integer M strictly less than M^{**} is not externally stable.)

Consequently, in order to show that there is no positive Nash equilibrium to the participation game, it is sufficient to show that $M^* < 1$. We now verify this inequality. The denominator of the right side of equation (38) is negative, so M^* is positive iff the numerator is also negative, i.e. iff

$$\frac{(\bar{p}+\theta_L)-2}{(\theta_L+\varepsilon-\bar{p})} < 1$$
$$\Longleftrightarrow$$
$$\bar{p} < 1 + \tfrac{\varepsilon}{2}.$$

When both the numerator and the denominator of equation (38) are negative, $M^* > 1$ iff

$$\frac{(\bar{p}+\theta_L)-2}{(\theta_L+\varepsilon-\bar{p})} < \frac{\bar{p}-\theta_L}{\varepsilon}.$$

This inequality is equivalent to

$$H\left(\bar{p}\right) \equiv \bar{p}^2 - 2\theta_L\bar{p} + \left(\theta_L - 2\right)\varepsilon + \left(\theta_L + \varepsilon\right)\theta_L < 0. \tag{39}$$

The function H is convex in $\bar{p}$ and it reaches a minimum at $\bar{p} = \theta_L$. Evaluating H at θ_L we have

$$H\left(\theta_L\right) = 2\left(\theta_L - 1\right)\varepsilon > 0.$$

Consequently, inequality (39) is never satisfied, so I conclude that for no parameter values is $M^* > 1$. Therefore, there are no positive equilibria to the participation game. ■

Proof. (**Lemma 1**) I use a proof by contradiction. Suppose that $M > \theta_L + \varepsilon$ in equilibrium. In this case it is optimal to require all countries to abate to capacity, and the price ceiling is never binding, i.e. there is no trade in equilibrium. From Section 2 it must be the case that

$$M = h\left(\theta_L + \frac{\varepsilon}{2}\right) < \theta_L + \frac{\varepsilon}{2} + 1.$$

For $\epsilon \geq 2$ we have

$$\theta_L + \frac{\varepsilon}{2} + 1 \leq \theta_L + \varepsilon$$

so

$$M < \theta_L + \varepsilon.$$

This inequality contradicts the hypothesis. ■

Proof. (Proposition 5) (i) Using equations (11) and (37) and the definition of δ, we have

$$\delta \equiv \min\left\{\frac{M-\theta_L}{\varepsilon}M, M\right\} - \min\left\{g\left(\frac{M-\theta_L}{\varepsilon}(M+1)\right), M\right\}. \tag{40}$$

When $\varepsilon \geq 2$, using Lemma 1, the first term of equation (41) is $\frac{M-\theta_L}{\varepsilon}M$. Now use the fact that M is an integer and the definition of the function $g(\cdot)$ to conclude

$$g\left(\tfrac{M-\theta_L}{\varepsilon}(M+1)\right) > M$$
$$\Longleftrightarrow$$
$$\tfrac{M-\theta_L}{\varepsilon}(M+1) \geq M+1.$$

The last inequality never holds by Lemma 1 and the assumption $\varepsilon \geq 2$, so the second term of equation (40) is $g\left(\frac{M-\theta_L}{\varepsilon}(M+1)\right)$. These two facts imply that equation (40) simplifies to

$$\delta = \frac{M-\theta_L}{\varepsilon}M - g\left(\frac{M-\theta_L}{\varepsilon}(M+1)\right) \tag{41}$$

when $\varepsilon \geq 2$. Using the definition of $g(\cdot)$,

$$\frac{M-\theta_L}{\varepsilon}(M+1) - 1 < g\left(\frac{M-\theta_L}{\varepsilon}(M+1)\right) \leq \frac{M-\theta_L}{\varepsilon}(M+1). \tag{42}$$

We have

$$\tfrac{M-\theta_L}{\varepsilon}M - \tfrac{M-\theta_L}{\varepsilon}(M+1) = -\tfrac{M-\theta_L}{\varepsilon} < 0$$

and

$$\tfrac{M-\theta_L}{\varepsilon}M - \left(\tfrac{M-\theta_L}{\varepsilon}(M+1) - 1\right) = 1 - \tfrac{M-\theta_L}{\varepsilon} > 0.$$

Using these two inequalities in inequality (42) produces inequality (21).

(ii) I use an example to show that δ can be positive negative: set $M = 6$, $\theta_L = 5$ and vary $\varepsilon \geq 2$. At all points of continuity, δ is decreasing in ε, and $\delta = 0$ at $\varepsilon = 6$ (a point of continuity). Therefore, δ can be positive or negative. Figure 5 shows the graph of δ. ■

Proof. (Proposition 6) Using Lemma 1 and equations (19) and (37), expected per country abatement under the price ceiling is

$$\frac{\bar{p} - \theta_L}{\varepsilon} < 1,$$

whereas equilibrium per country abatement in the absence of trade or the price ceiling is 1. ■

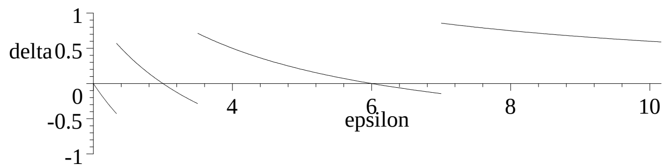

Figure 5: δ as a function of $\varepsilon \geq 2$ for $\theta_L = 5$ and $M = 6$

References

ALDY, J., A. KRUPNICK, R. NEWELL, I. PARRY, AND W. PIZER (2008): "Climate Economics and Policy," RFF Working Paper.

BARRETT, S. (1994): "Self-enforcing international environmental agreements," *Oxford Economic Papers*, 46, 878–894.

BARRETT, S. (2001): "International cooperation for sale," *European Economic Review*, 45, 1835–50.

——— (2002): "Consensus Treaties," *Journal of Institutional and Theoretical Politics*, 158, 519–41.

——— (2003): *Environment and Statecraft*. Oxford University Press.

BARRETT, S. (2006): "Climate Treaties and Breakthrough Technologies," *American Economic Review*, 96, 22–25.

BATABYAL, A. (2000): *The Economics of International Environmental Agreements*. Ashagate Press.

BLOCH, F. (1997): "Noncooperative models of coalition formation in games with spillovers," in *Dew Directions in the Economic Theory of the Environment*, ed. by C. Carraro, and D. Siniscalco, pp. 311– 352. Cambridge University Press.

BURGER, N., AND C. KOLSTAD (2009): "Voluntary public goods provision: coaltion formation and uncertainty," NBER Working Papers 15543.

CARRARO, C., AND D. SINISCALCO (1993): "Strategies for the International Protection of the Environment," *Journal of Public Economics*, 52, 309–328.

CHWE, M. S. (1994): "Farsighted Coalitional Stability," *Journal of Economic Theory*, 63, 299–325.

DANNENBERG, A., A. LANGE, AND B. STURM (2009): “On the formation of coalitions to provide public goods – experimental evidence from the lab,” Working paper.

DE ZEEUW, A. (2008): “Dynamic effects on the stability of international environmental agreements,” *Journal of Environmental Economics and Management*, 55(2), 163–74.

DIAMANTOUDI, E., AND E. SARTZETAKIS (2002): “International Environmental Agreements - The Role of Foresight,” University of Aarhus working paper 2002-10.

DIXIT, A., AND M. OLSON (2000): “Does Voluntary Participation undermine the Coase Theorem,” *Journal of Public Economics*, 76, 309 – 335.

EYCKMANS, J. (2001): “On the farsighted stability of the Kyoto Protocol,” CLIMNEG Working Paper 40, CORE, Universite Catholique de Louvain.

FINUS, M. (2001): *Game Theory and International Environmental Cooperation*. Edward Elgar.

FINUS, M., AND S. MAUS (2008): “Modesty may pay!,” *Journal of Public Economic Theory*, 10, 801–826.

FISCHER, C., AND R. MORGENSTERN (2006): “Carbon Abatement Costs: Why the Wide Range of Estimates?,” *The Energy Journal*, 27, 73–86.

FUJITA, T. (2006): “A comment on "International Cooperation for Sale",” *Economics Bulletin*, pp. 1–7.

HOEL, M. (1992): “International Environmental Conventions: the case of uniform reductions of emissions,” *Environmental and Resource Economics*, 2, 141–59.

KARP, L., AND J. ZHAO (2010): “A proposal for the design of the successor to the Kyoto Protocol,” in *Post-Kyoto International Climate Policy*, ed. by J. Aldy, and R. Stavins, pp. 530–562. Cambridge University Press, Cambridge.

KOLSTAD, C. D., AND A. ULPH (2008): “Uncertainty, learning and heterogeneity in international environmental agreements,” unpublished working paper.

KOSFELD, M., A. OKADA, AND A. RIEDL (2009): “Institution formation in public goods games,” *American Economic Review*, 99, 1335–1355.

MARIOTTI, M. (1997): “A Model of Agreements in Strategic Form Games,” *Journal of Economic Theory*, 74, 196–217.

McGinty, M. (2007): "International environmental agreements among asymmetric nations," *Oxford Economics Papers*, 59, 45 – 62.

Ray, D., and R. Vohra (2001): "Coalitional Power and Public Goods," *Journal of Political Economy*, 109(6), 1355 – 1382.

Xue, L. (1998): "Coalitional Stability under Perfect Foresight," *Economic Theory*, 11, 603–627.