Journal of Environmental Economics and Management 78 (2016) 67–84

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

International environmental agreements with consistent conjectures

Alejandro Gelves [a,b], Matthew McGinty [a,*]

[a] Department of Economics, University of Wisconsin-Milwaukee, PO Box 413, Milwaukee, WI 53201, United States
[b] University of North Texas, United States

ARTICLE INFO

Article history:
Received 14 August 2014
Available online 15 March 2016

JEL classification:
C7
D7
F5
H4
Q5

Keywords:
International environmental agreements
Consistent conjectures
Coalition formation
Public goods
Externalities

ABSTRACT

We introduce consistent conjectures into Barrett (1994) canonical model of international environmental agreements. The existing literature assumes inconsistent Nash conjectures, despite the fact that policymakers recognize that abatement levels are strategic substitutes and increases in abatement generate carbon leakage. With consistent conjectures much of the conventional wisdom is reversed. The non-cooperative abatement level is below the Nash equilibrium. The difference between Nash and consistent conjectures is greatest when benefits are large and costs are small. We find that large coalitions cannot form. However, small coalitions can result in substantial increases in abatement relative to the non-cooperative outcome.

Introduction

The current debate on International Environmental Agreements (IEAs) to reduce greenhouse gas emissions is dominated by the recognition of offsetting behavior. The fundamental issue is carbon leakage, which means that an increase in abatement by IEA signatories is met by a reduction in non-signatory abatement. Barrett (1994) is the seminal paper responsible for much of the conventional wisdom regarding IEAs. Barrett (1994), and all of the existing IEA literature, adopts Nash conjectures where each nation assumes that there is no response by other nations to changes in own abatement.

The Nash conjecture is a particularly bad assumption for nations' greenhouse gas emissions, as the entire concept of carbon leakage is an acknowledgement that other nations do respond. There is a large body of empirical literature on the effects of carbon leakage (for example Babiker, 2005; Elliott et al., 2010), yet the theoretical literature lags behind. No previous IEA has considered non-Nash conjectures and the resulting strategic implications of carbon leakage (see Finus, 2003 for a survey). This paper introduces conjectural variations to Barrett's model and builds an IEA that fits with the reality of carbon leakage as understood by actual policymakers.

While discussing the Kyoto Protocol in 1997 Senator Charles Hagel told the United States Senate the following (Kuik and Gerlagh, 2003). “The main effect of the assumed policy would be to redistribute output, employment and emissions from participating to non-participating countries.” Clearly, Senator Hagel anticipated carbon leakage and recognized that

* Corresponding author.
E-mail address: mmcginty@uwm.edu (M. McGinty).

http://dx.doi.org/10.1016/j.jeem.2016.02.005
0095-0696/

abatement levels are strategic substitutes. Subsequently, on July 25, 1997 the US Senate voted 95–0 in favor of the Byrd–Hagel resolution which stated that "The United States should not be a signatory to any protocol to, or other agreement regarding, the United Nations Framework Convention on Climate Change of 1992…"

More recently in September 2009 OECD Secretary General Gurria told the Informal Ministerial Meeting on Climate Change "If the EU were to cut emissions by 50% by 2050, with no other countries taking any action, our analysis suggests that almost 12% of their emissions reductions will be "leaked," or offset, through increased emissions in other countries."[1] This awareness also extends to the popular press and is part of the current public debate. In May 2014 Washington Post columnist Samuelson (2014) wrote "… any further U.S. emissions cuts would probably be offset by gains in China and elsewhere." Policymakers, negotiators and commentators are explicit in their non-Nash conjectures, however the existing IEA literature has adamantly retained a Nash conjecture of zero response.

This type of non-Nash behavior is called a conjectural variation in the oligopoly literature. With conjectural variations any change in own quantity is anticipated to induce a response by others (Bowley, 1924). Bresnahan (1981) extended this idea and proposed a consistent conjecture equilibrium (CCE). In a CCE the conjecture is assumed to be correct and equal the actual best-response slope.

The logical inconsistency of Nash conjectures is exacerbated when signatories to an IEA form coalitions. The IEA literature uses the concept of internal and external stability from D'Aspremont et al. (1983) for participation decisions, together with the assumption of Nash conjectures for abatement decisions. Internal stability means that no signatory would earn a higher payoff by leaving and external stability means that no non-signatory would earn a higher payoff by joining the agreement. A coalition of signatories is stable when it is both internally and externally stable. However, internal stability is itself by its very nature a non-Nash conjecture. Each nation compares their payoff as a coalition member with the potential payoff if they were to individually leave the coalition.

When a member considers leaving a coalition and reducing abatement it "correctly" anticipates (i) how the remaining signatories will reduce their abatement, (ii) how the other non-signatories will increase their abatement and (iii) how the nation that is leaving will best-respond to these changes. However, these "correctly" anticipated responses are determined by reaction functions that are derived from an inconsistent Nash conjecture. Thus, nations recognize the existence of carbon leakage when choosing to participate in an IEA, but do not recognize carbon leakage when choosing abatement levels.

Stackelberg models allow for a conjectural variation, and with a single follower the leader's conjecture is consistent and equal to the follower's best-response slope. When there are two or more followers the Stackelberg equilibrium with Nash conjectures is inconsistent since the conjecture determines the follower's aggregate best-response function. To the best of our knowledge, we are the first to explore this point. Barrett (1994) assumes that signatories to an IEA behave as a Stackelberg leader and collectively choose their most desired location on the aggregate best-response function of the followers. In Barrett (1994) the aggregate follower reaction function is obtained from Nash conjectures. Stackelberg leadership with more than one follower and Nash conjectures is itself an inconsistent equilibrium.

The use of consistent conjectures has been criticized as imposing an inherently dynamic process where players learn the best-response slopes of their rivals (Friedman, 1983, pp. 109–110). However, consistent conjectures have been applied to many situations where Nash conjectures are hard to justify. Consistent conjectures have been used in models of public goods (Cornes and Sandler, 1985), international trade (Fung, 1989), estimates of market power (Perloff et al., 2007), and strategic incentives in teams (Heywood and McGinty, 2012). In all of these areas player's recognize that their rivals' choice variables are either strategic substitutes or complements and do anticipate a response by others.

McGinty (2014) justifies the use of consistent conjectures in a two-nation version of Barrett (1994) model and provides the rationale for our approach. That paper shows that consistent conjectures emerge from individual payoff maximization in addition to the traditional approach of imposing consistency by assuming that conjectures match the actual best-response slope (Bresnahan, 1981; Kamien and Schwartz, 1983). There is a payoff advantage to recognizing offsetting behavior and that abatement levels are strategic substitutes. The CCE emerges as the unique subgame perfect equilibrium of a game where beliefs are chosen before abatement levels, as with actual policymakers. That paper shows that abatement is lower at the CCE than the Nash equilibrium (NE), a result first obtained for public goods by Sugden (1985). There is an individual payoff advantage to having a consistent conjecture, however both players are worse off at the CCE. Thus, there is a Prisoner's Dilemma in conjectures since the consistent conjecture dominates a Nash conjecture. The NE is not the appropriate non-cooperative outcome when policymakers have conjectures that differ from Nash. This paper shows how the difference in the non-cooperative outcome effects abatement and coalition stability in a model with n-nations.

The IEA is modeled as a three-stage game following Barrett (1994). In the first stage nations decide to participate in the agreement or not. In the second stage IEA signatories collectively choose abatement to maximize their combined payoff. In the third stage non-signatories individually chose abatement after observing signatory abatement. Barrett's central conclusion is that there is an inverse relationship between the gains to cooperation and the number of signatories to an IEA. Barrett finds that an IEA with full participation is possible, but only when there is essentially no difference between the Nash equilibrium of no cooperation and the social optimum of full participation.[2] This occurs when the benefits from abatement are large and the costs are low. However, this is precisely when the difference between the CCE and the NE is the largest.

[1] The full text is available at: http://www.oecd.org.reducinggreenhousgasemissionsindevelopedcountries.htm.

[2] Finus (2003) survey of the self-enforcing IEA literature indicates that large coalitions are only possible with both Stackelberg leadership and concave benefits.

We show that total abatement at the CCE is always below the NE level and that the difference is strictly increasing in the number of nations. With consistent conjectures signatories abate more than non-signatories. This result is in sharp contrast to Nash conjectures where small coalitions of signatories actually reduce abatement relative to the NE, a sort of reverse carbon leakage. Large coalitions cannot form when nations have consistent conjectures, contrary to the conventional Nash wisdom. Stable coalitions will not consist of more than three members. This is true even for parameter values where all 100 nations would be in a stable coalition with Nash conjectures (Barrett, 1994). However, small coalitions can close a large portion of the gap between the non-cooperative outcome and the optimum since total abatement at the CCE is so low. In contrast to Nash conjectures, our model generates carbon leakage as coalition members abate more than the non-cooperative outcome and non-members abate less. The main intuition is that with consistent conjectures the followers have a smaller incentive to decrease abatement than they would with Nash conjectures since each follower's incentive to reduce abatement is mitigated by the recognition that other followers also have an incentive to reduce. The free-rider problem is not simply that abatement is too low, but also the recognition that increases in abatement will result in carbon leakage.

Our results have practical implications for real-world IEA negotiations. If low levels of abatement (relative to the optimum) are observed in the absence of an agreement, one should not necessarily conclude that an IEA can achieve little, as the conventional wisdom dictates. Furthermore, consistent conjectures will never generate an IEA with full participation in Barrett (1994) model. Full participation can be obtained in a different model where abatement is a dominant strategy, but not when abatement levels are strategic substitutes.[3]

We consider three extensions to the benchmark model of a single coalition and identical nations. First, we allow for endogenous timing and show that leadership emerges as both the leader and the follower prefer the Stackelberg equilibrium to the simultaneous move game with consistent conjectures. This result differs from Nash conjectures where the simultaneous move game results with endogenous timing. With Nash conjectures the follower prefers the Nash equilibrium and the conventional wisdom of a simultaneous move game is obtained. Next, we allow for asymmetry and show how a mean-preserving spread improves abatement and payoffs in a two-nation model. We then allow for two-types of nations and show how the set of stable coalitions depends on an optimal transfer scheme. Finally, we allow for a “bottom-up” approach where more than one coalition can form via bilateral agreements. We show how this is compared to the single global agreement “top-down” approach found in most of the IEA literature.

The rest of the paper is organized as Section “Theoretical results” compares the Nash equilibrium, consistent conjectures equilibrium and social optimum. Section “Coalitions” allows for any coalition of signatories with the assumption of consistent conjectures. These results show how the coalition of signatories internalize the externalities among members and how the followers respond with consistent conjectures. Section “Simulation results” compares Barrett (1994) 100 nation results with those obtained by the CCE. Large coalitions cannot form, but small coalitions can close a large portion of the abatement gap between the non-cooperative and optimal outcomes. Section “Asymmetry” shows how asymmetry can slightly improve IEA performance and Section “Multiple coalitions” shows how membership can be increased if more than one coalition is allowed to form. The final section concludes and discusses directions for future research.

Theoretical results

Individual abatement is $q_i \geq 0$ and aggregate abatement is $Q = \sum_{i \in N} q_i$, where N is the set of nations with cardinality $n = |N|$. Individual payoffs are

$$\pi_i = \frac{b}{n}\left(aQ - \frac{Q^2}{2}\right) - \frac{c(q_i)^2}{2} \tag{1}$$

where a, b and c are strictly positive. Each nation has both identical benefit share $\frac{1}{n}$ and marginal abatement cost slope c. The first-order condition that maximizes individual payoff is

$$\frac{\partial \pi_i}{\partial q_i} = \frac{b}{n}\left[(a-Q)\left(1 + \sum_{j \neq i} r_{ij}\right)\right] - cq_i = 0 \tag{2}$$

where $r_{ij} \equiv \frac{\partial q_j}{\partial q_i}$ is nation i's conjecture about j's response. With Nash conjectures, nation i holds q_j constant thus $r_{ij} = 0$ $\forall i, j \in N$. In a conjectural variations model with strategic substitutes and offsetting behavior r_{ij} will be negative. This results in the best-response function

$$q_i^r = \frac{(a - Q_{-i})\left(1 + \sum_{j \neq i} r_{ij}\right)}{\gamma n + 1 + \sum_{j \neq i} r_{ij}} \tag{3}$$

where abatement by the other $n-1$ nations is denoted $Q_{-i} \equiv \sum_{j \neq i \in N} q_j$ and $\gamma \equiv \frac{c}{b}$ (Barrett, 1994). The Nash conjecture of $r = 0$

[3] Indeed, the discussion of the Montreal Protocol in the conclusion of Barrett (1994) indicates that US abatement of ozone depleting substances was a dominant strategy and would have occurred even without the agreement. Barrett cites EPA estimates that unilateral abatement of 50% by the US resulted in \$1373 billion of own benefit at a cost of only \$21 billion.

implies that the best-response slope is $\frac{-1}{\gamma n+1}$. With Nash conjectures as γ approaches zero the best-response slope approaches -1, or nearly complete offsetting behavior. Greenhouse gases mix uniformly in the upper atmosphere thus abatement is perfectly substitutable, and only the level of Q_{-i} matters, not the distribution. The consistent conjecture is the best-response slope.

$$r_{ji}=\frac{\partial q_i^r}{\partial q_j}=\frac{\partial q_i^r}{\partial Q_{-i}}=\frac{-\left(1+\sum_{j\neq i}r_{ij}\right)}{\gamma n+1+\sum_{j\neq i}r_{ij}} \tag{4}$$

Utilizing the symmetry $r_{ij}=r\ \forall i\neq j\in N$, the conjecture simplifies to

$$r=\frac{-(1+(n-1)r)}{\gamma n+1+(n-1)r} \tag{5}$$

which results in the quadratic

$$r^2(n-1)+rn(\gamma+1)+1=0 \tag{6}$$

The consistent conjecture is[4]

$$r=\frac{-n(\gamma+1)+\sqrt{n^2(\gamma+1)^2-4(n-1)}}{2(n-1)} \tag{7}$$

Since $r\in(-1,0)$ changes in abatement are partially offset by other nations. In the limit as γ approaches zero, r approaches $\frac{-1}{n+1}$, compared to the Nash conjecture best-response slope of -1. In one sense, this is the real free-rider problem. Not just the $\frac{1}{n}$ problem where each individual nation does not internalize the externality, but rather the recognition that increases in abatement will be offset by others, resulting in carbon leakage. In the limit as γ becomes large, r approaches 0. Next, putting the symmetric conjecture in the first-order condition (2) results in the CCE abatement level.

$$q^{cce}=\frac{a\left[2-n(\gamma+1)+\sqrt{n^2(\gamma+1)^2-4(n-1)}\right]}{n\left[(2-n)(\gamma+1)+\sqrt{n^2(\gamma+1)^2-4(n-1)}\right]} \tag{8}$$

Global abatement at the CCE is $Q^{cce}=nq^{cce}$.

$$Q^{cce}=\frac{a\left[2-n(\gamma+1)+\sqrt{n^2(\gamma+1)^2-4(n-1)}\right]}{(2-n)(\gamma+1)+\sqrt{n^2(\gamma+1)^2-4(n-1)}} \tag{9}$$

The aggregate abatement at the CCE is strictly decreasing in the number of nations.

$$\frac{\partial Q^{cce}}{\partial n}=\frac{-2a\gamma\left[(1+\gamma)\sqrt{n^2(1+\gamma)^2-4(n-1)}-n(1+\gamma)^2+2\right]}{\Delta}<0\ \forall n\geq 2\quad\text{and}\quad\gamma>0 \tag{10}$$

where $\Delta=\sqrt{n^2(1+\gamma)^2-4(n-1)}\left[(2-n)(1+\gamma)+\sqrt{n^2(1+\gamma)^2-4(n-1)}\right]^2>0\ \forall n\geq 2$ and $\gamma>0$. From Barrett (1994), Nash equilibrium abatement Q^{ne} is independent of the number of nations n.

$$Q^{ne}=\frac{a}{1+\gamma} \tag{11}$$

Therefore the abatement difference between the Nash equilibrium and the consistent conjectures equilibrium is increasing in the number of nations. In a Nash equilibrium the individual marginal benefit equals marginal cost and no nation has an incentive to change abatement since it believes that others will not change. By contrast, at the CCE individual marginal benefit is below marginal cost and when one nation considers increasing abatement it recognizes that this leads to carbon leakage as other nations reduce abatement. Together, these results show that carbon leakage is greatest for small γ when costs are low and benefits are high.

Aggregate Nash equilibrium abatement does not change in n, thus an important part of the $\frac{1}{n}$ problem is missing with Nash conjectures. This is an unusual model since the "number elasticity of individual Nash abatement" is -1. That is, if the number of nations increases by 1%, then Nash abatement for each nation decreases by 1%, leaving total abatement unchanged. From Barrett (1994) the optimal abatement levels $q^o$and $Q^o=nq^o$ maximize aggregate payoff and are

$$q^o=\frac{a}{n+\gamma}$$
$$Q^o=\frac{an}{n+\gamma} \tag{12}$$

[4] The negative root is not relevant since it would imply negative abatement levels by (2).

Clearly optimal abatement is increasing in n and decreasing in γ.

$$\frac{\partial Q^O}{\partial n} = \frac{a\gamma}{(n+\gamma)^2} > 0 \tag{13}$$

and of course the difference between the NE and optimum is increasing in n. Thus, the difference between the NE and the optimum is increasing in the number of nations. The difference between the CCE and the optimum is increasing in the number of nations not only because the optimum is increasing, but also because the CCE is decreasing in the number of nations.

Consistent conjectures recognizes an additional incentive to reduce abatement, magnifying the usual free-rider problem. With consistent conjectures reductions in abatement are partially offset by other nations' increase in abatement. Thus, there is a smaller reduction in benefit from reducing own abatement than with a Nash conjecture of no response. Since costs are always reduced by decreasing abatement, there is a greater incentive to reduce abatement with consistent conjectures. This incentive is greatest as the consistent best-response slope approaches -1, which occurs when costs are low and benefits are large. In this case own reductions in abatement are almost completely offset by others. The consistent conjecture free-rider problem is at its worst when the incentive to provide the public good is greatest.

For large γ there is little incentive to offset changes in abatement. This makes the consistent conjecture approach zero and the reaction function slope (for Nash conjectures) approaches zero. This is true regardless of n. However, there is a large difference between Q^{ne} and Q^{cce} when $\gamma = \frac{c}{b}$ is small since the difference in the best-response slopes is greatest. Barrett (1994) states "When c is small and b is large, unilateral abatement is substantial, and the gains to cooperation are relatively small." However, with consistent conjectures the opposite is true. Unilateral abatement is very small and the gains to cooperation are relatively large. Section "Simulation results" provides a complete comparison with Barrett's results.

Coalitions

We now consider stable coalitions and determine how much of the abatement gap between the non-cooperative outcome and the optimum can be obtained. In a Stackelberg model the three factors that influence the incentives to join coalitions are: (i) the timing advantage for coalition members, (ii) internalizing the positive externalities by coalition members, and (iii) cost-minimization when there is increasing marginal abatement cost by reallocating changes in abatement across coalition members. The timing effect increases the advantage to coalition membership, but as McGinty (2014) shows in a two-nation model, this effect alone leads to a reduction in abatement. The externality effect decreases the advantage to coalition membership. As the coalition increases in size the free-rider incentive becomes greater since each coalition member increases abatement to benefit other signatories. The n-nation model illustrates both the timing and externality effects and provides a basis for illustrating the role of consistent conjectures via a comparison with Barrett (1994). Let S be the set of leaders with cardinality $s = |S|$ being the number of signatories. Signatories choose abatement to maximize the sum of signatories payoffs called the coalition's worth. Signatories enjoy the timing advantage, internalize the positive externality among members and have consistent conjectures about both how the followers respond to each other and to the coalition's abatement.

Let T be the set of non-signatories with cardinality $t = |T|$ being the number of non-signatories. Each non-signatory maximizes individual payoff, but has consistent conjectures regarding the other non-signatories. Of course, all non-signatories have a consistent Nash conjecture with respect to the signatories $r_{ij} = 0\ \forall i \in T,\ \forall j \in S$. Each follower f maximizes

$$\pi_f = \frac{b}{n}\left(aQ - \frac{Q^2}{2}\right) - \frac{c\left(q_f\right)^2}{2} \tag{14}$$

The first-order condition is

$$\frac{\partial \pi_f}{\partial q_f} = \frac{b}{n}\left[(a-Q)\left(1 + \sum_{i \neq f \in T} r_{fi}\right)\right] - cq_f = 0 \tag{15}$$

The consistent conjecture for each follower is (derivation of what follows in Appendix A)

$$r_f = \frac{-(\gamma n + t) + \sqrt{d}}{2(t-1)} \tag{16}$$

where $d \equiv (\gamma n + t)^2 - 4(t-1)$. The individual and aggregate reaction function for the followers is

$$\begin{aligned} q_f &= (a - Q_s)\phi \\ Q_f &= t(a - Q_s)\phi \end{aligned} \tag{17}$$

where $\phi \equiv \left[\frac{2 - \gamma n - t + \sqrt{d}}{2\gamma n + t(2 - \gamma n - t + \sqrt{d})}\right]$.

The coalition's worth is

$$v(S) = \frac{sb}{n}\left[aQ - \frac{Q^2}{2}\right] - \frac{\sum_{j \in S} c\left(q_j\right)^2}{2} \tag{18}$$

Substituting in the followers' best-response and maximizing yield the individual and aggregate signatory abatement levels q_s.

$$q_s = \frac{as(1-t\phi)^2}{\gamma n + s^2(1-t\phi)^2}$$
$$Q_s = \frac{as^2(1-t\phi)^2}{\gamma n + s^2(1-t\phi)^2} \tag{19}$$

The follower abatement levels are

$$q_f = \frac{a\phi\gamma n}{\gamma n + s^2(1-t\phi)^2}$$
$$Q_f = \frac{at\phi\gamma n}{\gamma n + s^2(1-t\phi)^2} \tag{20}$$

Thus, global abatement is $Q = Q_s + Q_f$ where

$$Q = Q_s + Q_f = \frac{a\left[t\phi\gamma n + s^2(1-t\phi)^2\right]}{\gamma n + s^2(1-t\phi)^2} \tag{21}$$

Now that abatement levels for all possible coalition structures have been obtained, we can determine how the results depend on the parameters.

Simulation results

We first compare our results with the example in Barrett's (1994) Table 1. This example assumes 10 nations and shows abatement levels and payoffs for any number of signatories. We then compare our results with Barrett's for $n = 100$ and various values of b and c. From D'Aspremont et al. (1983) a coalition of size s is internally stable when no signatory has an incentive to leave, i.e. $\pi_f(s-1) \le \pi_s(s)$ and is externally stable when no non-signatory has an incentive to join, i.e. $\pi_f(s) < \pi_s(s+1)$. Thus, each nation has a non-Nash conjecture when considering participation decisions. With internal and external stability each nation recognizes that abatement changes in the number of signatories, however with Nash conjectures each nation does not recognize that other nations respond to changes in own abatement. We now show how consistent conjectures influences coalition stability, abatement and payoffs.

In Table 1 and Fig. 2 of Barrett (1994) the model is illustrated with an example where $n = 10, a = 100, b = 1, c = 0.25$. For this example the aggregate Nash equilibrium abatement is $Q^{ne} = 80.0$ with individual payoff $\pi^{NE} = 472.0$. We then take Table 1 and add values for consistent conjectures.

Following McGinty (2007) we present two measures of IEA performance. The first is the proportion of the abatement difference between the non-cooperative outcome and the optimum that is obtained by a stable IEA. The second,

Table 1
Stability analysis for consistent conjectures.

s	q_s		q_f		π_s		π_f		Q		Π	
	CC	NC	CC	NC	CC	NC	CC	NC	CC	NC	CC	NC
0	–	–	4.82	8.00	–	–	363.1	472.0	48.2	80.0	3631	4720
1	9.77	1.86	4.81	8.53	377.9	476.8	387.0	468.1	53.1	78.7	3861	4690
2	**15.24**	4.16	**4.14**	8.73	**404.8**	474.0	**431.6**	466.6	**63.6**	78.2	**4263**	4681
3	16.69	6.65	3.37	8.43	430.5	472.3	463.9	468.9	73.6	78.9	4538	4699
4	16.14	**8.91**	2.75	**7.57**	449.6	**472.2**	481.2	**474.9**	81.1	**81.1**	4685	**4738**
5	14.92	10.53	2.32	6.32	462.7	473.7	489.9	482.6	86.2	84.2	4763	4781
6	13.61	11.34	2.04	4.91	471.7	476.4	494.3	489.4	89.8	87.7	4807	4816
7	12.39	11.46	1.86	3.60	477.9	479.5	496.6	494.3	92.3	91.0	4835	4840
8	11.35	11.10	1.75	2.50	482.3	482.7	498.0	497.3	94.3	93.8	4854	4856
9	10.48	10.48	1.63	1.63	485.4	485.5	498.8	498.8	95.9	95.9	4868	4868
10	9.76	9.76	–	–	487.8	487.8	499.5	–	97.6	97.6	4878	4878

Numerical example with $n = 10, a = 100, b = 1, c = 0.25$. The stable IEA is given in bold. The consistent conjecture results are denoted CC and the Nash conjecture results from Barrett (1994) are denoted NC. Abatement by signatories, non-signatories and global abatement is denoted q_s, q_f and Q, respectively. Payoff for signatories, non-signatories and global payoff is denoted π_s, π_f and Π, respectively.

Table 2
IEA signatories.

b	c					
	0.01		1.00		100	
	CC	NC	CC	NC	CC	NC
0.01	3	3	3	2	2	2
1.00	2	51	3	3	3	2
100	1	100	2	51	3	3

The first number in each cell is the number of signatories with consistent conjectures (CC). The second shows Barrett (1994) results for Nash conjectures (NC). Assumes $a=1000$ and $n=100$.

effectiveness, normalizes this by the percentage difference between the optimum and non-cooperative outcomes. Effectiveness recognizes that closing a smaller proportion of a larger difference may be a more meaningful measure.

With no signatories, $s=0$, the NE is obtained with Nash conjectures (NC), while the CCE is obtained with consistent conjectures (CC). With full participation, $s=10$, the optimum is obtained regardless of conjectures. The coalition outcomes for $s=1$ to $s=9$ involve both Stackelberg leadership for the signatories and the type of Table 1 shows that the consistent conjecture results stand in stark contrast to Barrett (1994). With Nash conjectures, signatory abatement for the first three signatories is *below* the Nash equilibrium level $q^{ne}=8$. Thus, the first three signatories to an agreement actually reduce abatement, inducing the followers to increase their abatement. This strange result is a reverse form of carbon leakage where the agreement calls on the signatories to increase emissions. Barrett finds that the stable IEA consists of four members and global abatement is $Q=81.1$. The stable IEA results in closing 6% of the difference between the Nash equilibrium and optimal abatement $(\frac{Q(s)-Q^{ne}}{Q^{o}-Q^{ne}})$.

Aggregate abatement is only 48.2 at the CCE and $\pi^{cce}=363.1$. However, the first signatory to the IEA more than doubles abatement relative to $q^{cce}=4.82$. With CC, and at least two non-signatories, signatory abatement is always greater than non-signatory abatement. Abatement by non-signatories decreases as IEA membership increases, thus exhibiting carbon leakage. With consistent conjectures the stable IEA only consists of two members, but closes 31% of the much larger abatement gap $(\frac{Q(s)-Q^{cce}}{Q^{o}-Q^{cce}})$. For this example a coalition half as large can do five times as much. However, abatement under the stable IEA is still far below the Nash equilibrium. Rather, the Nash equilibrium is far too optimistic in terms of abatement when nations have consistent conjectures. The first signatory is starting from low abatement, thus the marginal net benefit is high. A high marginal payoff and a smaller amount of carbon leakage with consistent conjectures make the first leader increase abatement compared to the CCE. The CCE results in a much more realistic situation. There are no actual IEAs that dictate that the first signatories increase emissions as Nash conjectures imply.

Table 1 reveals another critical difference between CC and NC.[5] It has long been known in the industrial organization literature that it is difficult to obtain endogenous leadership. Gal-Or (1985) shows that when choice variables are strategic substitutes, as with public goods, the leader earns a higher payoff than the follower. Hamilton and Slutsky (1990) consider an extended game where players choose timing before their actions. The main result in Hamilton and Slutsky (1990) is to show that for endogenous leadership to emerge both the leader and the follower must prefer their payoff to that of the simultaneous move game. However, with Nash conjectures and strategic substitutes the follower prefers the Nash equilibrium and hence the simultaneous move game is the unique outcome with endogenous timing. Table 1 also illustrates this point for public goods. Nash equilibrium payoff is $\pi^{ne}=472$. With one signatory the leader has payoff $\pi_s=476.8>\pi^{ne}$, but each follower prefers the simultaneous move NE to the follower payoff of $\pi_f=468.1<\pi^{ne}$. Hence, with Nash conjectures and endogenous leadership the standard result for strategic substitutes is obtained. Endogenous leadership will not be an equilibrium.

To the best of our knowledge we are the first to consider endogenous timing with a consistent conjecture equilibrium in the simultaneous move game. Table 1 shows that payoff with CC at the simultaneous move outcome is $\pi^{cce}=363.1$. Consider the first signatory to an agreement. With CC both the leader and the followers prefer the payoffs of the sequential move game $\pi_s=377.9>\pi^{cce}$ and $\pi_f=387.0>\pi^{cce}$. The same is true when the second signatory joins the coalition as both leaders and followers earn a higher payoff. The reason for this difference is the effect mentioned above. With NC the first signatories exploit the timing advantage by actually reducing abatement, thus making followers worse off. By contrast CC becoming a signatory always implies increasing abatement, hence followers are also better off with leadership. This is an important result for understanding the current political landscape. With CC both climate leaders and followers have a higher payoff and prefer those roles to the simultaneous move outcome. This effect is not obtained with Nash conjectures where no nation wants to be a follower when the coalition is small.

We now turn to the $n=100$ nation results to show how the comparison depends on γ. Table 2 in Barrett (1994) shows the number of signatories out of 100 for various values of b and c. Again, we present his values and those for consistent conjectures.

[5] We would like to thank an anonymous referee and the Editor for suggesting this line of inquiry.

Table 3
Global abatement.

b	c					
	0.01		1.00		100	
	CC	NC	CC	NC	CC	NC
0.01	335.0	500.0	9.8	9.9	0.1	0.1
	357.4	503.9	10.4	10.1	0.1	0.1
	990.1		500.0		9.9	
	(3%)	(1%)	(0.12%)	(0.04%)	(0.02%)	(0.02%)
	(0.10)	(0.02)	(0.06)	(0.02)	(0.00)	(0.00)
1.00	502.5	990.1	335.0	500.0	9.8	9.9
	750.0	990.2	357.4	503.9	10.4	10.1
	999.9		990.1		500.0	
	(50%)	(1%)	(3%)	(1%)	(0.12%)	(0.04%)
	(0.99)	(0.01)	(0.10)	(0.02)	(0.06)	(0.02)
100	505.0	999.9	502.5	990.1	335.0	500.0
	980.6	1000.0	750.0	990.2	357.5	503.9
	1000.0		999.9		990.1	
	(96%)	(100%)	(50%)	(1%)	(3%)	(1%)
	(1.87)	(1.00)	(0.99)	(0.01)	(0.10)	(0.02)

The first rows in each cell are abatement for consistent conjectures (CC)Q^{CCE} and Nash conjectures (NC)Q^{NE}. The second row is abatement for the stable IEA, and the third is optimal abatement Q^o. The fourth row in each cell is the percentage of the abatement gap $\frac{Q(s)-Q^{ne}}{Q^o-Q^{ne}}$ or $\frac{Q(s)-Q^{cce}}{Q^o-Q^{cce}}$ closed by the IEA and the fifth is effectiveness $\left[\frac{Q(s)-Q^{ne}}{Q^o-Q^{ne}}\right]\frac{Q^o}{Q^{ne}}$ or $\left[\frac{Q(s)-Q^{cce}}{Q^o-Q^{cce}}\right]\frac{Q^o}{Q^{cce}}$. Note: $a=1000$ and $n=100$.

Table 4
Nation abatement.

b	c					
	0.01		1.00		100	
	CC	NC	CC	NC	CC	NC
0.01	3.35	5.00	0.10	0.10	0.00	0.00
	12.88	7.55	0.29	0.20	0.00	0.00
	3.29	4.96	0.10	0.10	0.00	0.00
	9.90		5.00		0.10	
1.00	5.02	9.90	3.35	5.00	0.10	0.10
248.73	10.00	12.88	7.55	0.29	0.20	
2.58	9.80	3.29	4.96	0.10	0.10	
10.00		9.90		5.00		
100	5.05	10.00	5.02	9.90	3.35	5.00
960.77	10.00	248.73	10.00	12.88	7.55	
0.20	–	2.58	9.80	3.29	4.96	
10.00		10.00		9.90		

The first row is abatement at the non-cooperative outcome (q^{cce} or q^{ne}), the second is signatory abatement (q_s), the third is non-signatory abatement (q_f) and the fourth is optimal abatement (q^o).

Result 1. With consistent conjectures (i) there are never more than three signatories to an IEA, and (ii) the difference in the number of signatories is greatest when benefits are high and costs are low.

Table 2 shows how consistent conjectures generates dramatically different results. When $c=0.01$ and $b=1$ there are only two signatories in a CCE compared to 51 in the NE. The most profound difference is when $c=0.01$ and $b=100$ where all 100 nations are signatories with NC and there is only one signatory with CC. For large values of $\gamma \equiv \frac{c}{b}$, there is almost no difference between the Nash equilibrium and the CCE. The consistent conjecture approaches the Nash conjecture zero for large γ. Consistent conjectures generate an even more dismal result in an already pessimistic literature. Abatement in the CCE is far below the Nash equilibrium when it is most important to form an agreement and the large coalitions, even full participation, that form with Nash conjectures are now only one or two members (small γ). Thus far, a more accurate conjecture leads to a much worse outcome. However, it is precisely when the outcome is so bad that the signatories can achieve the most. Table 3 provides the comparison.

Result 2. Under consistent conjectures, when benefits are high and costs are low coalitions can close a larger portion of the abatement gap. When benefits are low and costs are high a stable IEA results in little difference between the non-cooperative and IEA outcomes for both Nash and consistent conjectures.

Tables 2 and 3 show that when γ is very small there is a large difference between the CCE and NE. This means that a very small number of signatories can form a stable coalition that results in a very large difference with the non-cooperative outcome. Table 4 shows that signatories exploit their timing advantage by dramatically increasing abatement for small γ. With consistent conjectures the incentive for non-signatories to reduce abatement is offset by the other non-signatories, thus carbon leakage is reduced.

The results are profoundly different when γ is small. Global abatement at the Nash equilibrium when $\gamma=0.01$ is $Q^{ne}=990.1$. Table 2 shows the IEA with Nash conjectures consists of 51 nations, but the increase in abatement is trivial, 0.1. This is part of the conventional IEA wisdom: large coalitions can be stable but cannot achieve meaningful gains over the noncooperative outcome. However, this wisdom is overturned with consistent conjectures. Global abatement is only $Q^{cce}=502.5$, but a small coalition consisting of only two nations can overcome 50% of the abatement difference. In fact, it is precisely when the abatement difference is so large that IEAs with consistent conjectures can most improve on the non-cooperative outcome. The key is the timing advantage. The two signatories can provide abatement at such a low cost ($c=0.01$) that once that abatement is provided the followers incentives to reduce abatement are mitigated by each other. By contrast, with Nash conjectures the followers only respond to the leader and not each other.

Table 5 presents the payoffs.

Result 3. When benefits are high and costs are low a small number of signatories with consistent conjectures can close a large portion of both the abatement and payoff gaps, despite significant carbon leakage.

Tables 3–5 show that when γ is small, significant abatement is provided by a small number of signatories. Non-signatories reduce their abatement, resulting in significant carbon leakage. However, a self-enforcing IEA results in dramatic improvement over the non-cooperative outcome. Importantly, a large portion of the global payoff gap is closed by the actions of a small number of signatories. In situations where γ is small we should not expect the non-cooperative outcome to be the Nash equilibrium if policymakers have consistent conjectures. With consistent conjectures and small γ non-cooperative abatement is far below the Nash equilibrium. It is precisely in this case that we should expect a small number of signatories to provide a large increase in abatement, despite the resulting carbon leakage. All nations benefit and most or nearly all of the global payoff gap is closed. All nations earn a higher payoff than at the non-cooperative outcome. Obviously, signatories earn a lower payoff than non-signatories and this is the real impact of carbon leakage. Each nation would rather be a non-signatory, however since coalitions are internally stabile no IEA member has an incentive to leave.

Table 5
Global payoffs.

b	c					
	0.01		1.00		100	
	CC	Nash	CC	Nash	CC	Nash
0.01	2783 2928 4951 (7%) (0.12)	3738 3757 (2%) (0.02)	97 103 2500 (0.2%) (0.06)	98 100 (0.1%) (0.02)	1 1 50 (0%) (0.00)	1 1 (0%) (0.00)
1.00	376,228 468,129 499,950 (74%) (0.99)	499,902 499,903 (0.0%) (0.00)	278,329 292,789 495,050 (7%) (0.12)	373,750 375,659 (2%) (0.02)	9,709 10,277 250,000 (0.2%) (0.06)	9,803 9,990 (0.1%) (0.02)
100	37,749,926 49,976,528 49,999,950 (99.8%) (1.32)	49,999,949 49,999,950 (100%) (1.00)	37,622,782 46,812,891 49,995,000 (74%) (0.99)	49,990,197 49,990,293 (0%) (0.00)	27,832,911 29,278,934 49,504,950 (7%) (0.12)	37,375,000 37,565,851 (2%) (0.02)

The first rows in each cell are payoff for consistent conjectures (CC) and Nash conjectures (NC). The second row is payoffs for the stable IEA, and the third is optimal payoffs. The fourth row in each cell is the percentage of the payoff gap closed by the IEA and the fifth is effectiveness. Note: $a=1000$ and $n=100$.

We extend the basic model of a single coalition of identical nations from Barrett (1994) in two additional directions. First, we examine the impact of benefit and cost asymmetry with CC. Second, we allow for more than one coalition. That is, a “bottom-up” approach where multiple bilateral agreements may form, rather than a single “top-down” global agreement. Both of these extensions result in substantial increase in the complexity of the problem, hence we are limited to numerical rather than analytical results for some of what follows.

Asymmetry

We relax the assumption of symmetry and show how the consistent conjectures equilibrium changes with a mean-preserving spread. We do this for a two-nation model to isolate the role of asymmetric benefit shares and marginal abatement cost slopes. We then allow for n nations consisting of two types and show the impact of asymmetry on coalition formation and provision.

Two asymmetric nations

We first consider two asymmetric nations with benefit shares α_i and α_j where $\alpha_i+\alpha_j=1$. With asymmetric costs the marginal abatement cost curves have slope c_i. The payoff function in Eq. (1) becomes

$$\pi_i=\alpha_i b\left(aQ-\frac{Q^2}{2}\right)-\frac{c_i(q_i)^2}{2}. \tag{22}$$

Repeating the procedure in Section “Theoretical results” yields the best-response functions

$$\begin{aligned} q_i^r&=\frac{\left(a-q_j\right)(1+r_{ij})}{\theta_i+1+r_{ij}}\\ q_j^r&=\frac{(a-q_i)(1+r_{ji})}{\theta_j+1+r_{ji}} \end{aligned} \tag{23}$$

where $\theta_i\equiv\frac{c_i}{\alpha_i b}$ and with symmetry $\theta_i=\gamma n$. Thus, the consistent conjectures are

$$\begin{aligned} r_{ji}&=\frac{\partial q_i^r}{\partial q_j}=\frac{-(1+r_{ij})}{\theta_i+1+r_{ij}}\\ r_{ij}&=\frac{\partial q_j^r}{\partial q_i}=\frac{-(1+r_{ji})}{\theta_j+1+r_{ji}}. \end{aligned} \tag{24}$$

To simplify what follows, note that

$$\begin{aligned} r_{ji}+1&=\frac{\theta_i}{\theta_i+1+r_{ij}}\\ r_{ij}+1&=\frac{\theta_j}{\theta_j+1+r_{ji}}. \end{aligned} \tag{25}$$

This is more complicated than the symmetric result from Section “Theoretical results” since the conjectures are now a simultaneous system of distinct quadratic equations. Substituting the second equation into the first implies (25) results in the quadratic (derivation of what follows in Appendix B)

$$(r_{ji})^2\theta_i+r_{ji}(\theta_i+\theta_j+\theta_i\theta_j)+\theta_j=0. \tag{26}$$

Repeating this procedure for the other conjecture and recognizing that the positive root is relevant we obtain the solutions[6]

$$\begin{aligned} r_{ji}&=\frac{-(\theta_i+\theta_j+\theta_i\theta_j)+\sqrt{(\theta_i+\theta_j+\theta_i\theta_j)^2-4\theta_i\theta_j}}{2\theta_i}=\frac{\mu}{2\theta_i}\\ r_{ij}&=\frac{-(\theta_i+\theta_j+\theta_i\theta_j)+\sqrt{(\theta_i+\theta_j+\theta_i\theta_j)^2-4\theta_i\theta_j}}{2\theta_j}=\frac{\mu}{2\theta_j} \end{aligned} \tag{27}$$

[6] Note that with symmetry $\theta_i=\theta_j=\theta=\gamma n=\frac{2c}{b}$ and (27) reduces to the two-nation version of (7), $r=\frac{-(2+\theta)+\sqrt{\theta(4+\theta)}}{2}$. For example, if c=0.25 and b=1 then both (27) and (7) show $r=-0.5$ and if c=1 and b=1 then $r=-0.27$.

Table 6
Two-nation asymmetry.

	Symmetry	Identical θ	Positive covariance	Negative covariance
	$\alpha_i=\alpha_j=\frac{1}{2}$ $c_i=c_j=1$ $\theta_i=\theta_j=2$	$\alpha_i=\frac{1}{4}\ \alpha_j=\frac{3}{4}$ $c_i=0.5\ c_j=1.5$ $\theta_i=\theta_j=2$	$\alpha_i=\frac{1}{3}\ \alpha_j=\frac{2}{3}$ $c_i=0.5\ c_j=1.5$ $\theta_i=1.5\ \theta_j=2.25$	$\alpha_i=\frac{2}{3}\ \alpha_j=\frac{1}{3}$ $c_i=0.75\ c_j=1.25$ $\theta_i=1.125\ \theta_j=3.75$
ρ	−0.58	−0.58	−0.55	−0.52
q_i^{cce}	21.13	21.13	28.50	40.18
q_j^{cce}	21.13	21.13	16.21	7.82
Q^{cce}	42.26	42.26	44.71	48.00
π_i^{cce}	1443.38	721.69	954.17	1826.67
π_j^{cce}	1443.38	2165.06	2117.30	2386.12
Π^{cce}	2886.75	2886.75	3071.48	4212.78

The values in the table are for $a=100$ and $b=1$, where $\theta_i \equiv \frac{c_i}{b\alpha_i}$.

where $\mu \equiv -(\theta_i+\theta_j+\theta_i\theta_j)+\sqrt{(\theta_i+\theta_j+\theta_i\theta_j)^2-4\theta_i\theta_j}$. This clearly implies that the relative conjectures equals the relative θ's.

$$\frac{r_{ji}}{r_{ij}}=\frac{\theta_j}{\theta_i}=\frac{\left(\frac{c_j}{\alpha_j}\right)}{\left(\frac{c_i}{\alpha_i}\right)} \tag{28}$$

So, nation j's conjecture about i's response, r_{ji} is small when c_i is large and α_i is small such that $r_{ji}\left(\frac{c_i}{\alpha_i}\right)=r_{ij}\left(\frac{c_j}{\alpha_j}\right)$ holds. Thus, there is a small degree of carbon leakage by the high cost, low benefit nation. The simultaneous best-response from (23), allowing for any conjectures, simplifies to

$$q_i=\frac{a\theta_j(1+r_{ij})}{\theta_i(1+r_{ji})+\theta_j(1+r_{ij})+\theta_i\theta_j}$$
$$q_j=\frac{a\theta_i(1+r_{ji})}{\theta_i(1+r_{ji})+\theta_j(1+r_{ij})+\theta_i\theta_j}. \tag{29}$$

Again, note that (27) implies that $1+r_{ij}=\frac{\mu+2\theta_j}{2\theta_j}$ and $1+r_{ji}=\frac{\mu+2\theta_i}{2\theta_i}$. Then returning the conjectures (27) to the simultaneous best-responses in (29) results in

$$q_i^{cce}=\frac{a(\mu+2\theta_j)}{2\sqrt{(\theta_i+\theta_j+\theta_i\theta_j)^2-4\theta_i\theta_j}}.$$

Similarly, abatement by nation j at the CCE is

$$q_j^{cce}=\frac{a(\mu+2\theta_i)}{2\sqrt{(\theta_i+\theta_j+\theta_i\theta_j)^2-4\theta_i\theta_j}}. \tag{30}$$

Global abatement at the asymmetric CCE is $Q^{cce}=q_i^{cce}+q_j^{cce}$ (derivation in Appendix B)

$$Q^{cce}=a(\rho+1)$$

where the key ratio is $\rho\equiv\frac{-\theta_i\theta_j}{\sqrt{(\theta_i+\theta_j+\theta_i\theta_j)^2-4\theta_i\theta_j}}<0$. Thus, Q^{cce} increases as ρ approaches zero from below.[7] Now it becomes clear how mean-preserving asymmetry affects abatement.

Result 4. Aggregate abatement at the consistent conjectures equilibrium is minimized for identical nations or when the marginal abatement cost slope-benefit share ratio is identical ($\theta_i=\theta_j$). Aggregate abatement exceeds the symmetric level for both a negative and a positive covariance between costs and benefits.

Proof in Appendix C.

Table 6 compares symmetry to asymmetry with a positive and a negative covariance between MAC slopes and benefit shares. Note that identical θ's can occur with a positive covariance, even when nations are asymmetric. Both types of asymmetry result in greater abatement and payoff than would be obtained with identical θ's.

Eq. (10) indicates that abatement at the CCE is declining in the number of nations, hence to understand the impact of asymmetry we need to go beyond a two-nation model. To repeat this procedure for large n would involve n first order

[7] Note that the root in the denominator of ρ is strictly positive since $\sqrt{(\theta_i+\theta_j+\theta_i\theta_j)^2-4\theta_i\theta_j}=\sqrt{(\theta_i-\theta_j)^2+\theta_i\theta_j(\theta_i\theta_j+2\theta_i+2\theta_j)}>0$.

Table 7
Asymmetric abatement.

$s1$	$s2$					
	0	1	2	3	4	5
0	– – $q_f^1=5.8$ $q_f^2=3.8$	– $q_s^2=8.3$ $q_f^1=5.8$ $q_f^2=3.8$	– $q_s^2=13.3$ $q_f^1=5.0$ $q_f^2=3.3$	– $q_s^2=15.1$ $q_f^1=4.1$ $q_f^2=2.7$	– $q_s^2=14.9$ $q_f^1=3.4$ $q_f^2=2.2$	– $q_s^2=14.0$ $q_f^1=2.8$ –
1	$q_s^1=10.8$ $q_f^1=5.9$ $q_f^2=3.8$	$q_s^1=12.6$ $q_s^2=18.8$ $q_f^1=4.9$ $q_f^2=3.2$	$q_s^1=12.4$ $q_s^2=18.6$ $q_f^1=4.0$ $q_f^2=2.6$	$q_s^1=11.5$ $q_s^2=17.2$ $q_f^1=3.2$ $q_f^2=2.2$	$q_s^1=10.4$ $q_s^2=15.6$ $q_f^1=2.7$ $q_f^2=1.8$	$q_s^1=9.4$ $q_s^2=14.0$ $q_f^1=2.3$ –
2	$q_s^1=16.5$ – $q_f^1=5.0$ $q_f^2=3.3$	$q_s^1=14.9$ $q_s^2=22.4$ $q_f^1=4.0$ $q_f^2=2.6$	$q_s^1=13.1$ $q_s^2=19.7$ $q_f^1=3.2$ $q_f^2=2.1$	$q_s^1=11.5$ $q_s^2=17.2$ $q_f^1=2.7$ $q_f^2=1.8$	$q_s^1=10.1$ $q_s^2=15.2$ $q_f^1=2.3$ $q_f^2=1.6$	$q_s^1=9.0$ $q_s^2=13.5$ $q_f^1=2.1$ –
3	$q_s^1=17.8$ – $q_f^1=4.1$ $q_f^2=2.7$	$q_s^1=14.9$ $q_s^2=22.4$ $q_f^1=3.2$ $q_f^2=2.2$	$q_s^1=12.7$ $q_s^2=19.0$ $q_f^1=2.7$ $q_f^2=1.8$	$q_s^1=11.0$ $q_s^2=16.5$ $q_f^1=2.3$ $q_f^2=1.6$	$q_s^1=9.6$ $q_s^2=14.5$ $q_f^1=2.1$ $q_f^2=1.4$	$q_s^1=8.6$ $q_s^2=12.9$ $q_f^1=2.0$ –
4	$q_s^1=17.0$ – $q_f^1=3.4$ $q_f^2=2.2$	$q_s^1=14.0$ $q_s^2=21.1$ $q_f^1=2.8$ $q_f^2=1.8$	$q_s^1=11.9$ $q_s^2=17.9$ $q_f^1=2.4$ $q_f^2=1.6$	$q_s^1=10.3$ $q_s^2=15.5$ $q_f^1=2.1$ $q_f^2=1.4$	$q_s^1=9.1$ $q_s^2=13.7$ $q_f^1=2.0$ $q_f^2=1.3$	$q_s^1=8.2$ $q_s^2=12.3$ $q_f^1=1.8$ –
5	$q_s^1=15.6$ – – $q_f^2=1.9$	$q_s^1=13.0$ $q_s^2=19.4$ – $q_f^2=1.6$	$q_s^1=11.1$ $q_s^2=16.6$ – $q_f^2=1.5$	$q_s^1=9.7$ $q_s^2=14.6$ – $q_f^2=1.4$	$q_s^1=8.6$ $q_s^2=13.0$ – $q_f^2=1.2$	$q_s^1=7.8$ $q_s^2=11.7$ – –

conditions, each of which is a function of conjectures regarding the other $n-1$ nations. The conjectures are themselves a system of n quadratic equations. Then this system would need to be simultaneously solved, each solution of which would then result in a distinct quadratic which is then a function of $n-1$ distinct parameter ratios. We are unable to obtain a general solution to the roots these quadratics. Even if we limit ourselves to two types of n nations we are only able to solve this system numerically. However, this is typical of the IEA literature. Barrett (1994) resorts to simulations to find coalition stability in a much simpler model with identical nations and Nash conjectures of zero.

Two types of nations

Next, we allow for an asymmetric version of Table 1 which assumes $c=0.25$ and $\alpha=\frac{1}{n}=\frac{1}{10}$. To illustrate the effect of asymmetry on coalition stability we take Table 1 and consider a mean-preserving spread of the benefit and cost parameters. To compare with the 10-nation results we have 5 type 1 nations and 5 type 2 nations. The number of type 1 signatories is $s1$ and type 2 signatories is $s2$. Type 1 nations are high benefit and high cost slope with $c_1=0.3$ and $\alpha_1=\frac{2}{15}$ and type 2 nations have benefit $\alpha_2=\frac{1}{15}$ and cost slope $c_2=0.2$. The remaining parameters are $a=100$ and $b=1$.

Table 7 shows the abatement levels for signatories (q_s^i) and non-signatories (q_f^i) for the possible coalition structures.

Table 8 shows the payoffs for each coalition structure and the aggregate payoff, or worth (V), of each coalition. In the absence of transfers a coalition is internally stable when no signatory would earn a higher payoff from the resulting coalition structure if they were to leave. A coalition is externally stable if no non-signatory would earn a higher payoff by joining the coalition. Stable coalitions are those that are both internally and externally stable for both types of nations.

Internal and external stability for type 1 nations is determined via each column. For example, begin with no agreement at $s1=s2=0$ and hold $s2$ at zero. Since $\pi_f^1=483<\pi_s^1=505$ at $s1=1$ the first type 1 nation joins the coalition. Since $\pi_f^1=517$ at $s1=1$ is less than $\pi_s^1=543$ at $s1=2$ the second type 1 nation joins the coalition. However, the free-rider payoff $\pi_f^1=580$ at $s1=2$ is greater than the signatory payoff $\pi_s^1=578$ at $s1=3$ so the third type 1 nation has no incentive to join. Hence, $s1=2$ is the unique internally and externally stable coalition structure for the type 1 nations when $s2=0$. Applying this analysis to the other columns in Table 8 reveals the other stable coalitions for the type 1 nations. Stable coalitions are similarly determined via each row for type 2 nations.

A stable coalition must be internally and externally stable for both types of nations. The result is a unique internally and externally stable coalition consisting of two type 1 nations and no type 2 nation with payoffs $\pi_s^1=543$ and $\pi_f^2=291$. Global abatement and payoff at the stable coalition are $Q=64.7$ and $\Pi=4279$. In the absence of transfers there is the same number

Table 8
Coalition stability without transfers.

$s1$	$s2$: 0	1	2	3	4	5
0	–	$\pi_s^2=251$	$\pi_s^2=267$	$\pi_s^2=\mathbf{283}$	$\pi_s^2=296$	$\pi_s^2=305$
	$\pi_f^1=483$	$\pi_f^1=510$	$\pi_f^1=565$	$\pi_f^1=609$	$\pi_f^1=635$	$\pi_f^1=\mathbf{649}$
	$\pi_f^2=242$	$\pi_f^2=256$	$\pi_f^2=\mathbf{283}$	$\pi_f^2=305$	$\pi_f^2=318$	–
1	$\pi_s^1=505$	$\pi_s^1=557$	$\pi_s^1=597$	$\pi_s^1=\mathbf{621}$	$\pi_s^1=\mathbf{636}$	$\pi_s^1=645$
	–	$\pi_s^2=\mathbf{255}$	$\pi_s^2=275$	$\pi_s^2=291$	$\pi_s^2=302$	$\pi_s^2=310$
	$\pi_f^1=517$	$\pi_f^1=577$	$\pi_f^1=617$	$\pi_f^1=640$	$\pi_f^1=651$	$\pi_f^1=658$
	$\pi_f^2=260$	$\pi_f^2=289$	$\pi_f^2=309$	$\pi_f^2=320$	$\pi_f^2=326$	–
2	$\pi_s^1=\mathbf{543}$	$\pi_s^1=\mathbf{590}$	$\pi_s^1=\mathbf{618}$	$\pi_s^1=634$	$\pi_s^1=644$	$\pi_s^1=650$
	–	$\pi_s^2=262$	$\pi_s^2=283$	$\pi_s^2=297$	$\pi_s^2=307$	$\pi_s^2=313$
	$\pi_f^1=580$	$\pi_f^1=622$	$\pi_f^1=643$	$\pi_f^1=653$	$\pi_f^1=659$	$\pi_f^1=662$
	$\pi_f^2=\mathbf{291}$	$\pi_f^2=311$	$\pi_f^2=322$	$\pi_f^2=327$	$\pi_f^2=329$	–
3	$\pi_s^1=578$	$\pi_s^1=612$	$\pi_s^1=631$	$\pi_s^1=642$	$\pi_s^1=649$	$\pi_s^1=653$
	–	$\pi_s^2=273$	$\pi_s^2=291$	$\pi_s^2=303$	$\pi_s^2=310$	$\pi_s^2=316$
	$\pi_f^1=623$	$\pi_f^1=644$	$\pi_f^1=654$	$\pi_f^1=659$	$\pi_f^1=662$	$\pi_f^1=664$
	$\pi_f^2=\mathbf{312}$	$\pi_f^2=322$	$\pi_f^2=327$	$\pi_f^2=330$	$\pi_f^2=331$	–
4	$\pi_s^1=603$	$\pi_s^1=626$	$\pi_s^1=639$	$\pi_s^1=647$	$\pi_s^1=652$	$\pi_s^1=656$
	–	$\pi_s^2=284$	$\pi_s^2=298$	$\pi_s^2=308$	$\pi_s^2=314$	$\pi_s^2=318$
	$\pi_f^1=645$	$\pi_f^1=655$	$\pi_f^1=660$	$\pi_f^1=663$	$\pi_f^1=664$	$\pi_f^1=665$
	$\pi_f^2=\mathbf{323}$	$\pi_f^2=328$	$\pi_f^2=330$	$\pi_f^2=331$	$\pi_f^2=332$	–
5	$\pi_s^1=620$	$\pi_s^1=636$	$\pi_s^1=645$	$\pi_s^1=651$	$\pi_s^1=655$	$\pi_s^1=657$
	–	$\pi_s^2=293$	$\pi_s^2=304$	$\pi_s^2=311$	$\pi_s^2=316$	$\pi_s^2=319$
	–	–	–	–	–	–
	$\pi_f^2=\mathbf{328}$	$\pi_f^2=330$	$\pi_f^2=332$	$\pi_f^2=332$	$\pi_f^2=333$	–

The number of type 1 (2) signatories is s_1 (s_2). The aggregate payoff, or worth, of each coalition is denoted V. The payoff for a type i signatory is π_s^i and a type i non-signatory is π_f^i. Stable coalitions for each type are in bold. Note: $n=10, a=100, b=1$.

Table 9
Coalition stability with transfers.

$s1$	$s2$: 0	1	2	3	4	5
0		$V=251$	$V=534$	$V=849$	$V=1184$	$V=1526$
		$\sigma=9$	$\sigma=22$	$\sigma=\mathbf{0}$	$\sigma=-36$	$\sigma=-64$
1	$V=505$	$V=812$	$V=1147$	$V=1494$	$V=1844$	$V=2193$
	$\sigma=22$	$\sigma=87$	$\sigma=\mathbf{4}$	$\sigma=-42$	$\sigma=-71$	$\sigma=-86$
2	$V=1086$	$V=1443$	$V=1803$	$V=2161$	$V=2515$	$V=2865$
	$\sigma=\mathbf{51}$	$\sigma=-2$	$\sigma=-53$	$\sigma=-85$	$\sigma=-95$	$\sigma=-96$
3	$V=1733$	$V=2110$	$V=2476$	$V=2835$	$V=3189$	$V=3538$
	$\sigma=-7$	$\sigma=-68$	$\sigma=-97$	$\sigma=-105$	$\sigma=-108$	$\sigma=-103$
4	$V=2412$	$V=2789$	$V=3155$	$V=3512$	$V=3863$	$V=4211$
	$\sigma=-80$	$\sigma=-110$	$\sigma=-117$	$\sigma=-114$	$\sigma=-109$	$\sigma=-105$
5	$V=3100$	$V=3472$	$V=3834$	$V=4188$	$V=4537$	$V=4883$
	$\sigma=-125$	$\sigma=-131$	$\sigma=-126$	$\sigma=-123$	$\sigma=-111$	$\sigma=-107$

Stable coalitions (i.e. those with a positive surplus σ and whose enlargement results in a negative surplus) are in bold. The worth (aggregate payoff) to each coalition is denoted by V. Note: $n=10, a=100, b=1, c=0.25$.

of signatories as in Table 1 with symmetric nations where abatement and payoff at the stable coalition are $Q=63.6$ and $\Pi=4263$. Asymmetry alone does very little to improve the outcome in the absence of transfers.

Table 8 also illustrates how endogenous leadership will emerge with asymmetric nations. Both type 1 and type 2 followers prefer those roles to payoff at the simultaneous move game with CC. Consider the first type 1 signatory with no type 2 signatories, $s1=1, s2=0$. Type 1 followers have a higher payoff $\pi_f^1=517$ than at the CCE $\pi_f^1=483$. Similarly, the type 1 leader has a higher payoff $\pi_s^1=505$ than at the CCE $\pi_f^1=483$ so all type 1 nations prefer endogenous leadership. The same is true for the type 2 followers who also have a higher payoff as a follower $\pi_f^2=260$ than at the CCE $\pi_f^2=242$. The same analysis holds for the first type 2 signatory, or the second type 1 signatory, as all nations have a higher payoff with endogenous leadership.

Table 10
Two coalitions.

Coalitions	q_s	Q_s	q_f	Q	π_s	π_f	Π	π_f^{S-i}
$s_1 = s_2 = 0$	–	–	4.82	48.2	–	363.1	3631	–
$s_1 = s_2 = 1$	18.31	8.31	4.97	56.4	396.1	401.7	4006	387.0
$\mathbf{s_1 = s_2 = 2}$	**10.12**	**20.24**	**4.63**	**68.2**	**436.8**	**446.9**	**4428**	**424.7**
$s_1 = s_2 = 3$	19.52	28.57	4.76	76.2	460.3	468.8	4637	474.2
$s_1 = s_2 = 4$	8.70	34.82	5.77	81.2	472.8	478.1	4739	486.4
$s_1 = s_2 = 5$	8.44	42.19	–	84.4	478.9	–	4789	490.2

Note: $n = 10, a = 100, b = 1, c = 0.25$. Stable coalition structure is in bold.

There is a substantial and rapidly growing literature on optimal transfer schemes in IEAs (Carraro et al., 2006; McGinty, 2007; Weikard, 2009; McGinty, 2011). A credible transfer requires asymmetry since both nations must be better off post-transfer. It has been well established in this literature that the set of stable coalitions without transfers is a proper subset of the set of stable coalitions with transfers (Pavlova and De Zeeuw, 2013; Finus and McGinty, 2015). That is, a properly designed transfer scheme can achieve any outcome that would be obtained without transfers, and potentially much better outcomes. With an optimal transfer scheme a coalition is internally stable when the worth exceeds the sum of the payoffs that each member would earn if they were to individually leave the coalition. This difference is called the surplus. If the surplus is positive then the worth is sufficient to deter any member from free-riding. One simple transfer scheme from McGinty (2011) awards each member their outside payoff plus an equal share of the surplus. This rule has been shown to be the most robust since it minimizes the incentives for deviations across all nations by equating the incentive to do so (Table 9).

With optimal transfers the set of stable coalitions is the edge between those coalitions with a positive surplus whose enlargement of either type would lead to a negative surplus. Hence, the set of (weakly) stable coalitions is $\{(s1 = 2, s2 = 0), (s1 = 1, s2 = 2), (s1 = 0, s2 = 3)\}$. The results are not remarkably different with two types of asymmetric nations. Global abatement and payoff at the three stable coalitions are $\{(Q = 64.7, \Pi = 4279), (Q = 73.5, \Pi = 4544), (Q = 71.3, \Pi = 4505)\}$, respectively. Abatement and payoff are higher at the three member coalitions, but remain far below the optimum in Table 1 (Table 10).

Multiple coalitions

Next, we allow for more than one coalition. Previous work has shown that multiple small coalitions may be able to achieve more than one large coalition (Carraro and Marchiori, 2003; Finus and Rundshagen, 2003; Carraro and Büchner, 2005). Carraro and Büchner (2005) refer to this as a "bottom-up" approach where multiple bilateral negotiations occur, rather than the "top-down" approach of Kyoto where there is a single agreement for all nations. They consider a two-bloc coalition structure, with those outside the agreement behaving as free-riders. Using the FEEM-RICE model they show that two coalitions, each of which with two members, can be a stable coalition structure. With multiple coalitions members only internalize the positive externality within, not across, coalitions. This implies a smaller increase in abatement from joining when there are multiple small coalitions compared to a single large coalition.

Following this approach, we take Table 1 and allow for two coalitions. The coalitions choose simultaneously with respect to each other and each follower observes the coalitions' abatement then individually chooses abatement. Each coalition has consistent conjectures with respect to both the followers and with respect to the other coalition. A coalition recognizes the direct effect from its own abatement on followers and the indirect effect on followers from its influence on the other coalitions' abatement. Hence, the coalitions behave as a Stackelberg leaders with respect to the followers and have consistent conjectures with respect to each other. The aggregate best-response of the followers is unchanged from Eq. (41) in Appendix A. We consider coalition stability for two equally sized coalitions and compare the results with Table 1.

The payoff that each member would get if they were to leave is π^f_{S-i}. To determine internal stability each member compares their signatory payoff with what they would receive if they were to individually leave the coalition. For example, if the coalition structure is $s_1 = s_2 = 4$ with two non-signatories, then if a member were to leave the resulting coalition structure would be one coalition with three members, one coalition with four members and three non-signatories. The payoff from leaving a four member coalition is then $\pi_f^{S-i} = 486.4$. The two member coalitions are internally stable since $\pi_s = 436.8$ exceeds $\pi_f^{S-i} = 424.7$ and they are externally stable since the three member coalitions have $\pi_f^{S-i} = 474.2$ which exceeds $\pi_s = 460.3$. The stable coalition structure is $s_1 = s_2 = 2$, with aggregate abatement $Q = 68.2$. Thus, allowing for two coalitions results in a slight improvement over the single stable coalition result with two members and aggregate abatement $Q=63.6$.

Conclusion

All previous IEAs have adopted Nash conjectures, even in models where abatement levels are strategic substitutes. The best-response functions imply carbon leakage, however with Nash conjectures this incentive is ignored. We show that consistent conjectures reverses much of the conventional wisdom regarding IEAs. Consistent conjectures generate an understanding of IEAs that account for carbon leakage and the free-rider problem. It is not just that the public goods externality is not internalized, but also that increases in abatement by signatories will be met be reductions by non-signatories.

When benefits are high and costs are low there is very little difference between the Nash equilibrium and global optimum and IEAs with full participation are stable. We show that abatement is always lower in a CCE and that the difference with the Nash equilibrium is increasing in the number of nations. In the benchmark model we find that stable coalitions cannot be larger than three, even in situations where Nash conjectures result in stable coalitions of all 100 nations. Specifically, our results differ dramatically when the benefits from abatement are high and the costs are low, the most relevant situation for real-world policy since this is when the incentive to provide public goods is the greatest. In this case we find that the CCE is far below the NE. However, small coalitions can overcome a substantial amount of the difference between the CCE and the optimum, despite carbon leakage. Carbon leakage does not occur with Nash conjectures where the first few signatories to an IEA actually reduce abatement. The policy implications are that small coalitions can result in a substantial improvement over the non-cooperative outcome when carbon leakage is more clearly understood via consistent conjectures.

With endogenous timing Stackelberg leadership will emerge with consistent conjectures, while the simultaneous move game is obtained with Nash conjectures. We show that asymmetry or multiple coalitions can slightly increase IEA membership and abatement, but that the results are still substantially below the Nash equilibrium and the social optimum.

Our results point to several directions for future research. Investigating the link between consistent conjectures and endogenous leadership could provide many new insights to models of industrial organization and mergers. The recent growth of the experimental IEA literature provides another direction for future research. Do experimental subjects learn the best-response slope and anticipate carbon leakage, or do they have naive Nash conjectures? Do subjects converge to the CCE as models of learning or evolutionary models would predict? The role of group size and timing are also potential treatment variables. Finally, it would be interesting to investigate consistent conjectures in IEAs with different benefit and cost functions.

Acknowledgment

We would like to thank Dan Friedman and John Heywood for numerous discussions and insights regarding both the topic and this paper. We are grateful for the constructive comments by seminar participants at University of Aberdeen and University of Richmond. The paper has also benefited greatly from the comments of two anonymous referees and the Editors. The usual caveat applies.

Appendix A

The set of non-signatories is T with cardinality $t=|T|$. Each follower f maximizes

$$\pi_f = \frac{b}{n}\left(aQ - \frac{Q^2}{2}\right) - \frac{c\left(q_f\right)^2}{2}. \tag{31}$$

The first-order condition is

$$\frac{\partial \pi_f}{\partial q_f} = \frac{b}{n}\left[(a-Q)\left(1+\sum_{i \neq f \in T} r_{fi}\right)\right] - cq_f = 0. \tag{32}$$

Each follower has the reaction function

$$q_f^r = \frac{(a-Q_{-f})\left(1+\sum_{i \neq f \in T} r_{fi}\right)}{\gamma n + 1 + \sum_{i \neq f \in T} r_{fi}}. \tag{33}$$

The slope of any given follower's reaction function is

$$r_{if} \equiv \frac{\partial q_f^r}{\partial q_{-f}} = \frac{\partial q_f^r}{\partial Q_{-f}} = \frac{-\left(1+\sum_{i \neq f \in T} r_{fi}\right)}{\gamma n + 1 + \sum_{i \neq f \in T} r_{fi}}. \tag{34}$$

Recognizing the symmetry, $r_{fi} = r_{if} \equiv r_f$ and given that there are t followers $\sum_{i \neq f \in T} r_{fi} = (t-1)r_f$. Thus we have

$$r_f = \frac{-(1+(t-1)r_f)}{\gamma n + 1 + (t-1)r_f} \tag{35}$$

which results in the quadratic

$$(r_f)^2(t-1)+r_f(\gamma n+t)+1=0. \tag{36}$$

Note that this is the same as the quadratic and consistent conjecture found earlier when $t=n$. The consistent conjecture for each follower is then

$$r_f=\frac{-(\gamma n+t)+\sqrt{d}}{2(t-1)} \tag{37}$$

where $d\equiv(\gamma n+t)^2-4(t-1)$. Putting the consistent conjecture into the first-order condition for each follower we have

$$\gamma n q_f=(a-Q)\left[1+(t-1)r_f\right]. \tag{38}$$

Which simplifies to

$$\gamma n q_f=(a-Q)\left[\frac{2-\gamma n-t+\sqrt{d}}{2}\right]. \tag{39}$$

Recognizing that $Q=tq_f+Q_S$ this becomes

$$\gamma n q_f=\left(a-Q_s-tq_f\right)\left[\frac{2-\gamma n-t+\sqrt{d}}{2}\right]q_f\left[\gamma n+t\left(\frac{2-\gamma n-t+\sqrt{d}}{2}\right)\right]=(a-Q_s)\left[\frac{2-\gamma n-t+\sqrt{d}}{2}\right]. \tag{40}$$

The reaction function for individual and aggregate followers is

$$q_f=(a-Q_s)\left[\frac{2-\gamma n-t+\sqrt{d}}{2\gamma n+t\left(2-\gamma n-t+\sqrt{d}\right)}\right] \tag{41}$$

$$q_f=(a-Q_s)\phi$$
$$Q_f=t(a-Q_s)\phi$$

where $\phi\equiv\left[\frac{2-\gamma n-t+\sqrt{d}}{2\gamma n+t\left(2-\gamma n-t+\sqrt{d}\right)}\right]$. The coalition's worth is

$$\upsilon(S)=\frac{sb}{n}\left[aQ-\frac{Q^2}{2}\right]-\frac{\sum_{j\in S}c\left(q_j\right)^2}{2}. \tag{42}$$

Substituting in the best-response of the followers we have

$$\upsilon(S)=\frac{sb}{n}\left[a\left(\sum_{j\in S}q_j+t\phi\left(a-\sum_{j\in S}q_j\right)\right)-\frac{\left[\sum_{j\in S}q_j+t\phi\left(a-\sum_{j\in S}q_j\right)\right]^2}{2}\right]-\frac{\sum_{j\in S}c\left(q_j\right)^2}{2}. \tag{43}$$

The first-order condition is

$$\frac{\partial\upsilon(S)}{\partial q_j}=\frac{sb}{n}\left[a(1-t\phi)-(1-t\phi)\left[\sum_{j\in S}q_j+t\phi\left(a-\sum_{j\in S}q_j\right)\right]\right]-cq_j=0. \tag{44}$$

Given the symmetry each signatory has an identical first-order condition thus, $\sum_{j\in S}q_j=sq_s$ and we have

$$\frac{\gamma n q_s}{s}=(1-t\phi)\left[a-sq_s-t\phi\left(a-sq_s\right)\right]\frac{\gamma n q_s}{s}=(1-t\phi)^2\left[a-sq_s\right]q_s\left[\frac{\gamma n}{s}+s(1-t\phi)^2\right]=a(1-t\phi)^2q_s=\frac{as(1-t\phi)^2}{\gamma n+s^2(1-t\phi)^2}. \tag{45}$$

The aggregate abatement by the leaders is $Q_s=sq_s$ which is

$$Q_s=\frac{as^2(1-t\phi)^2}{\gamma n+s^2(1-t\phi)^2}. \tag{46}$$

The follower abatement levels are

$$Q_f=t(a-Q_s)\phi$$
$$Q_f=\frac{at\phi\gamma n}{\gamma n+s^2(1-t\phi)^2}. \tag{47}$$

Each individual follower has abatement level

$$q_f=\frac{a\phi\gamma n}{\gamma n+s^2(1-t\phi)^2}. \tag{48}$$

Global abatement is $Q=Q_s+Q_f$

$$Q = Q_s + Q_f = \frac{a\left[t\phi\gamma n + s^2(1-t\phi)^2\right]}{\gamma n + s^2(1-t\phi)^2}. \tag{49}$$

Appendix B

Eq. (25) results in the quadratic

$$r_{ji} = \frac{-\left(\dfrac{\theta_j}{\theta_j+1+r_{ji}}\right)}{\left[\theta_i + \dfrac{\theta_j}{\theta_j+1+r_{ji}}\right]} \tag{50}$$

$$r_{ji} = \left(\frac{-\theta_j}{\theta_j+1+r_{ji}}\right)\left(\frac{\theta_j+1+r_{ji}}{\theta_i\left[\theta_j+1+r_{ji}\right]+\theta_j}\right)$$

$$r_{ji}\left(\theta_i\left[\theta_j+1+r_{ji}\right]+\theta_j\right)+\theta_j = 0$$

$$\left(r_{ji}\right)^2\theta_i + r_{ji}\left(\theta_i+\theta_j+\theta_i\theta_j\right)+\theta_j = 0.$$

Repeating this for the other conjecture and solving the quadratics result in

$$r_{ji} = \frac{-\left(\theta_i+\theta_j+\theta_i\theta_j\right)+\sqrt{\left(\theta_i+\theta_j+\theta_i\theta_j\right)^2-4\theta_i\theta_j}}{2\theta_i} = \frac{\mu}{2\theta_i} \tag{51}$$

$$r_{ij} = \frac{-\left(\theta_i+\theta_j+\theta_i\theta_j\right)+\sqrt{\left(\theta_i+\theta_j+\theta_i\theta_j\right)^2-4\theta_i\theta_j}}{2\theta_j} = \frac{\mu}{2\theta_j}$$

where $\mu \equiv -\left(\theta_i+\theta_j+\theta_i\theta_j\right)+\sqrt{\left(\theta_i+\theta_j+\theta_i\theta_j\right)^2-4\theta_i\theta_j}$. Hence, the relative conjectures equal the relative θ's.

$$\frac{r_{ji}}{r_{ij}} = \frac{\theta_j}{\theta_i} = \frac{\left(\dfrac{c_j}{\alpha_j}\right)}{\left(\dfrac{c_i}{\alpha_i}\right)} \tag{52}$$

The simultaneous best-response to (23) is

$$q_i = \frac{a\theta_j(1+r_{ij})}{\theta_i(1+r_{ji})+\theta_j(1+r_{ij})+\theta_i\theta_j}\; q_j = \frac{a\theta_i(1+r_{ji})}{\theta_i(1+r_{ji})+\theta_j(1+r_{ij})+\theta_i\theta_j}. \tag{53}$$

Substituting the conjectures (27) into (29) results in

$$q_i = \frac{a\theta_j\left[\dfrac{\mu+2\theta_j}{2\theta_j}\right]}{\theta_i\left[\dfrac{\mu+2\theta_i}{2\theta_i}\right]+\theta_j\left[\dfrac{\mu+2\theta_j}{2\theta_j}\right]+\theta_i\theta_j} \tag{54}$$

$$q_i = \frac{a\left[\dfrac{\mu+2\theta_j}{2}\right]}{\left[\dfrac{\mu+2\theta_i}{2}\right]+\left[\dfrac{\mu+2\theta_j}{2}\right]+\theta_i\theta_j}$$

$$q_i = \frac{\dfrac{a}{2}\left(\mu+2\theta_j\right)}{\mu+\theta_i+\theta_j+\theta_i\theta_j}$$

$$q_i^{cce} = \frac{a\left(\mu+2\theta_j\right)}{2\sqrt{\left(\theta_i+\theta_j+\theta_i\theta_j\right)^2-4\theta_i\theta_j}}.$$

Similarly, for j we have

$$q_j^{cce} = \frac{a(\mu+2\theta_i)}{2\sqrt{\left(\theta_i+\theta_j+\theta_i\theta_j\right)^2-4\theta_i\theta_j}}. \tag{55}$$

Global abatement is

$$Q^{cce} = q_i^{cce} + q_j^{cce} = \frac{a\left(\mu + 2\theta_j + \mu + 2\theta_i\right)}{2\sqrt{\left(\theta_i + \theta_j + \theta_i\theta_j\right)^2 - 4\theta_i\theta_j}}\, Q^{cce} = \frac{a\left(\mu + \theta_i + \theta_j\right)}{\sqrt{\left(\theta_i + \theta_j + \theta_i\theta_j\right)^2 - 4\theta_i\theta_j}}$$

$$Q^{cce} = \frac{a\left(-\theta_i\theta_j + \sqrt{\left(\theta_i + \theta_j + \theta_i\theta_j\right)^2 - 4\theta_i\theta_j}\right)}{\sqrt{\left(\theta_i + \theta_j + \theta_i\theta_j\right)^2 - 4\theta_i\theta_j}}\, Q^{cce} = a\left[\frac{-\theta_i\theta_j}{\sqrt{\left(\theta_i + \theta_j + \theta_i\theta_j\right)^2 - 4\theta_i\theta_j}} + 1\right] Q^{cce} = a(\rho + 1)$$

where $\rho \equiv \frac{-\theta_i\theta_j}{\sqrt{\left(\theta_i + \theta_j + \theta_i\theta_j\right)^2 - 4\theta_i\theta_j}} < 0.$

Appendix C

Without loss of generality let $\theta_i \geq \theta_j > 0$ where $\theta_i \equiv \frac{c_i}{b\alpha_i}$. Define the difference $d \equiv \theta_i - \theta_j \geq 0$ and the mean as $m = \frac{\theta_i + \theta_j}{2}$. We can then write the θ's in terms of the mean and difference, $\theta_i = m + \frac{d}{2}$ and $\theta_j = m - \frac{d}{2}$. The upper bound on the difference, given the mean, is then $\theta_j = m - \frac{d}{2} > 0$ or $d < 2m$. We can then write the key ratio $\rho \equiv \frac{-\theta_i\theta_j}{\sqrt{\left(\theta_i + \theta_j + \theta_i\theta_j\right)^2 - 4\theta_i\theta_j}}$ in terms of m and d to obtain

$$\rho = \frac{d^2 - 4m^2}{\sqrt{16\left(m^4 + d^2 - md^2\right) + 64m^3 - 8m^2d^2 + d^4}} \tag{56}$$

Taking the derivative with respect to the difference d and setting it equal to zero result in three roots, but only one is relevant $d^* = 0$. The other two roots are $d = \frac{\pm 2m\sqrt{(m-1)(m+1)}}{m-1}$. The roots are imaginary for $m < 1$ and the negative root is not in the relevant range since $d < 0$. The positive root is also not in the relevant range since $d = \frac{2m\sqrt{(m-1)(m+1)}}{m-1} > 2m$ for $m > 1$ which implies $\theta_j < 0$. Thus, we have a single real root at $d^* = 0$. The second-order condition evaluated at $d^* = 0$ is $\frac{\partial^2\rho}{\partial d^2} = \frac{m+1}{m(m+4)\sqrt{m^3(m+4)}} > 0$, hence ρ and Q^{cce} are minimized for identical θ's.

References

Babiker, M., 2005. Climate change policy, market structure, and carbon leakage. J. Int. Econ. 65, 421–445.
Barrett, S., 1994. Self-enforcing international environmental agreements. Oxf. Econ. Pap. 46, 878–894.
Bresnahan, Timothy F., 1981. Duopoly models with consistent conjectures. Am. Econ. Rev. 71 (5), 934–945.
Bowley, A., 1924. Mathematical Groundwork of Economics. Oxford University Press, Oxford, UK.
Carraro, C., Büchner, B., 2007. Regional and Sub-global Climate Blocs. A Game-theoretic Perspective on Bottom-up Climate Regimes. Working Paper No. 21.2005. Fondazione Eni Enrico Mattei (FEEM).
Carraro, C., Marchiori, C., 2003. Stable coalitions in endogenous formation of economic coalitions. Edward Elgar, Cheltenham, UK156–198.
Carraro, C., Eyckmans, J., Finus, M., 2006. Optimal transfers and participation decisions in international environmental agreements. Rev. Int. Org. 1 (4), 379–396.
Cornes, R., Sandler, T., 1985. On the consistency of conjectures with public goods. J. Publ. Econ. 27, 125–129.
D'Aspremont, C., Jacquemin, A., Gabszewicz, J.J., Weymark, J., 1983. On the stability of collusive price leadership. Can. J. Econ. 16, 17–25.
Elliott, J., Foster, I., Kortum, S., Munson, T., Cervantes, F.P., Weisbach, D., 2010. Trade and carbon taxes. Am. Econ. Rev. Pap. Proc. 100 (2), 465–469.
Finus, M., 2003. Stability and design of international environmental agreements: the case of transboundary pollution. In: Folmer, H., Tietenberg, T. (Eds.), International Yearbook of Environmental and Resource Economics, 2003/4, Edward Elgar, Cheltenham, UK, pp. 82–158.
Finus, M., Rundshagen, B., 2003. Endogenous coalition formation in global pollution control: a partition function approach. In: Carraro, C. (Ed.), Endogenous Formation of Economic Coalitions, Edward Elgar, Cheltenham, UK, pp. 199–243.
Finus, M., McGinty, M., 2015. The Anti-Paradox of Cooperation: Diversity Pays! Bath Economics Research Working Papers 40/15. University of Bath.
Friedman, J., 1983. Oligopoly Theory. Cambridge University Press, New York, USA.
Fung, K.C. 1989. Tariffs, quotas, and international oligopoly. Oxf. Econ. Pap. 41, 749–757.
Gal-Or, E., 1985. First mover and second mover advantages. Int. Econ. Rev. 26, 649–652.
Hamilton, J., Slutsky, S., 1990. Endogenous timing in duopoly games: Stackelberg or Cournot equilibria. Games Econ. Behav. 2, 29–46.
Heywood, J., McGinty, M., 2012. Scale economies, consistent conjectures and teams. Econ. Lett. 117, 566–568.
Kamien, M., Schwartz, N., 1983. Conjectural variations. Can. J. Econ. 16, 191–211.
Kuik, O., Gerlagh, R., 2003. Trade liberalization and carbon leakage. Energy J. 24 (3), 97–120.
McGinty, M., 2007. International environmental agreements among asymmetric nations. Oxf. Econ. Pap. 59 (1), 45–62.
McGinty, M., 2011. A risk-dominant allocation: maximizing coalition stability. J. Publ. Econ. Theory 13 (2), 311–325.
McGinty, M., 2014. Rational Conjectures in Public Good Games. Working Paper. Department of Economics, University of Wisconsin-Milwaukee.
Pavlova, Y., De Zeeuw, A., 2013. Asymmetries in international environmental agreements. Environ. Dev. Econ. 18, 51–68.
Perloff, J., Karp, L., Golan, A., 2007. Estimating Market Power and Strategies. Cambridge University Press, New York, USA.
Samuelson, Robert., 2014. On Climate Change, We Have No Solution. The Washington Post, May 12, 2014.
Sugden, R., 1985. Consistent conjectures and voluntary contributions to public goods: why the conventional theory does not work. J. Publ. Econ. 27, 117–124.
Weikard, H.P., 2009. Cartel stability under optimal sharing rule. Manch. Sch. 77 (5), 575–593.