Appalachian
STATE UNIVERSITY®
BOONE, NORTH CAROLINA

Endogenous Minimum Participation In International Environmental Agreements: An Experimental Analysis

By: **David M. McEvoy**, **Todd L. Cherry**, & John K. Stranlund

Abstract

Almost all international environmental treaties require a minimum number of countries to ratify the treaty before it enters into force. Despite the wide-spread use of this mechanism, little is known about its effectiveness at facilitating cooperation. We analyze an agreement formation game that includes an endogenously determined minimum participation constraint and then test the predictions using economic experiments. We demonstrate theoretically that players will vote to implement an efficient coalition size as the membership requirement and this coalition will form. Experimental tests of the theory demonstrate that the minimum participation mechanism is highly effective at facilitating cooperation when efficiency requires the participation of all players. However, when efficiency requires only a subset of players to participate, profitable coalitions are often deliberately blocked. In light of our results it is possible that equity concerns can impede the formation of international agreements when membership requirements allow free riders.

1 Introduction

Almost all international environmental agreements (IEAs) require a minimum number of countries to ratify the agreement before it enters into force. If the minimum membership condition is met, as well as other accompanying conditions, the treaty becomes active and ratifying countries are bound to their commitments. If the condition is not met, the treaty never enters into force and affected countries are not bound by its provisions. Minimum membership requirements are standard in treaties that address global environmental issues, and Barrett (2003, pp. 165–194) provides a thorough examination of membership requirements in international environmental agreements. For example, the Kyoto Protocol to the United Nations Framework Convention on Climate Change required ratification by at least 55 parties prior to its entry into force in February 2005. Similarly, the Montreal Protocol on Substances that Deplete the Ozone Layer required at least 11 countries to ratify it before it entered into force in 1989.[1] Minimum membership requirements extend to treaties that address international risk and security, such as the Treaty on the Non-Proliferation of Nuclear Weapons, which required the ratification of the five nuclear nations (at that time) plus 40 additional nations. The Chemical Weapons Convention required ratification by at least 65 nations before it entered into force in 1997. While many treaties require ratification by a subset of affected parties to enter into force, some require all parties to join. For example, the Convention for the Protection of the Marine Environment of the North-East Atlantic required accession of all contracting parties, and the Treaty of Lisbon required ratification by all European Union member states.

Because of their wide-spread use in international agreements, there is a small game-theoretic literature on the role of minimum membership requirements. Black et al. (1992) find that an exogenously imposed minimum membership requirement can increase participation with an international environmental agreement. A similar finding is discussed in Carraro and Siniscalco (1992). Clearly, however, national sovereignty requires that all provisions of international treaties, including membership requirements, be determined endogenously by affected parties. Carraro et al. (2009) provide an analysis of endogenous minimum participation requirements in IEAs, and find that the mechanism increases the size of cooperative coalitions relative to the non-cooperative baseline. Particularly relevant for our study, Carraro et al. also find that endogenous minimum participation requirements can lead to efficient outcomes. [2]

However, empirical data on the effectiveness of minimum participation constraints in IEAs is absent because of the lack of counterfactuals and controls necessary for careful analysis. Consequently, we use laboratory experiments to evaluate the effectiveness of endogenously determined minimum requirements in motivating coalition formation to provide public goods. Of course, laboratory experiments, like the game-theoretic models they stem from, dramatically simplify the international negotiation process. However, the experimental approach allows us to examine the performance of particular institutions in a controlled environment. Since international environmental agreements are designed to address collective action prob-

---

[1] Both the Kyoto and Montreal protocols included an extra provision to ensure participants represented a minimum level of global emissions. Members to the Kyoto Protocol had to represent at least 55 % of the total 1990 greenhouse gas emissions. Similarly, the Montreal Protocol required representation of at least two-thirds of the total 1986 consumption of ozone-depleting substances.

[2] A recent working paper by Weikard et al. (2009) extends the theoretical analysis by Carraro et al. (2009) to include heterogeneous agents under different sharing rules. Harstad (2006) models coalitions of heterogeneous agents that contribute to a public good and derives optimal participation rules as well as solving for political equilibria (i.e., Condorcet winners).

lems between resource users, the same fundamental tensions that exist between countries can be simulated in a laboratory environment, and the results of experiments can shed light on how institutional rules, in this case minimum participation rules, affect coalition formation.[3]

We begin by developing a theoretical model of endogenous agreement formation to provide a public good, including the endogenous determination of a minimum membership requirement. Our model is a simplified version of the one in Carraro et al., yielding clearer predictions so that it is amenable to tests with experimental data. The game we analyze consists of two stages. The first stage is unique in the experimental coalition formation literature in that individuals vote on a minimum membership requirement. In the second stage, each individual decides whether they will join a coalition to provide the public good and a coalition forms if the membership requirement is met. We demonstrate that expected-payoff-maximizing individuals will choose to adopt the minimum membership requirement that is equal to the efficient coalition size, and that this coalition will form. This theoretical result is robust to whether the efficient agreement is the grand coalition or a smaller coalition.

We designed our laboratory experiments to test these predictions. Results show that when efficiency requires all players to join an agreement, subjects adopted the efficient minimum membership requirement about 75 % of the time. Conditional on subjects adopting the efficient minimum membership requirement, efficient agreements formed over 90 % of the time. Consequently, in these cases we observe significantly higher efficiency than other coalition formation experiments that do not include endogenous minimum membership requirements (e.g., Kosfeld et al. 2009; Dannenberg et al. 2010; McEvoy et al. 2011). The results were not as positive when efficiency required only 50 % participation. While subjects still adopted the efficient minimum membership requirement about 75 % of the time, the efficient coalition formed just over half of the time. Interestingly, the efficient coalition was blocked in about one-third of the cases in which it was adopted as the membership condition. While this is inconsistent with a theoretical model of individuals with standard preferences, we demonstrate that this behavior is consistent with a theory that allows for preferences concerning inequality in the manner of Fehr and Schmidt (1999).[4] The lesson for international environmental policy making is clear: endogenous minimum participation requirements are effective at achieving high levels of cooperation, but they may be significantly less effective when the participation requirements allow for freeriders.

Although no experimental research has investigated the effectiveness of minimum membership requirements directly, a significant literature has done so indirectly with analyses of *minimum contributing sets* in public good games (e.g., Rapoport 1985; Erev and Rapoport 1990; Cadsby and Maynes 1999; Cooper and Stockman 2002; McEvoy 2010). In these games, all individual contributions to the public good are returned if there are too few contributors. In contrast with our game, minimum contributing sets are not determined by the players themselves. A related literature on endogenous institution formation in public goods games has received much attention recently. Most of this literature explores the formation of institutions that govern all players (e.g., Walker et al. 2000; Gurerk et al. 2006; Tyran and Feld 2006; Kroll et al. 2007; Sutter et al. 2010). Typically, however, international treaties are formed to restrict the behavior of only participating nations. A handful of recent studies that are motivated by international treaty formation examine coalition formation in public goods

---

[3] The use of laboratory experiments in the evaluation of public policies is well established (Plott 1987; Shogren and Hurley 1999; Cason and Plott 1996; Stranlund et al. 2011), and experiments are particularly well suited to evaluate the effectiveness of different voluntary institutions.

[4] Others have conjectured that preferences over inequality might play important roles in international cooperation because all treaties must tackle issues of equity (Ringius et al. 2002; Lange and Vogt 2003).

games (e.g., Kosfeld et al. 2009; Dannenberg et al. 2010; McEvoy et al. 2011; Barrett and Dannenberg 2012), but none consider endogenous minimum membership requirements.[5]

The remainder of the paper proceeds as follows. In the next section we develop our theoretical model and derive its predictions. In Sect. 3 we describe the experimental design we use to test these predictions. We present the results of the experiments in Sect. 4 and conclude in Sect. 5.

2 Theory

Consider a game in which n players decide whether to contribute a single unit to a public good. Our model is a variant of the seminal model by Barrett (1994). Let the number of individuals who contribute to the public good (and the total supply of the public good) be s. Let $b > 0$ denote the shared benefit players receive from contributions to the public good up to $\bar{s} \leq n$ and let $c > 0$ denote the individual cost of contributing. The basic payoff function of a player that contributes to the public good is

$$\pi_i = A + bs - c \text{ for } s \leq \bar{s}, \tag{1}$$

where A is a positive constant. Assume that $c > b$ so that no player would contribute to the public good in a standard, non-cooperative Nash equilibrium. However, suppose that collective welfare is maximized when $\bar{s}$ individuals contribute their units to the public good. In the typical coalition formation game, collective welfare is maximized when everyone contributes; that is, when $\bar{s} = n$. However, we also allow for the possibility that it is efficient that the public good be provided by a coalition that is smaller than the grand coalition. We assume a simple case of this in which individual contributions yield a return of b up to an aggregate level and additional contributions yield a return of zero. For example, abatement activities that move emissions levels below the absorptive capacity of the natural environment will not yield additional environmental benefits. Examples of this also include any project for which contributions in excess of what is required to provide the good do not yield public benefits. In either case, if $\bar{s}$ is efficient, then $n(A + b\bar{s}) - c\bar{s} > nA$, which requires $nb > c$.

Given the motivation for collective action to provide the public good, suppose that the players are able to form a cooperative coalition in a two-stage game. In the first stage (the *voting stage*), all players vote on the minimum number of members required for a coalition to form. Players cannot vote for a membership requirement equal to zero and they cannot opt out of this vote. We call the outcome of this vote the *minimum membership requirement*. In our experiments we implement a plurality voting rule so that the number that receives the most votes becomes the membership requirement.[6]

---

[5] The sequence of decisions in some of the coalition formation experiments differs significantly from the sequence of decisions in international treaty formation. For example, in the experiments of Kosfeld et al. (2009), players first decide whether to join a coalition, and then the members vote on whether to contribute to the public good. Hence, in their analysis members of a coalition decide what the coalition should accomplish *after* they make their participation decision. In contrast, the players in our study understand ex ante what they are required to do in a coalition before they make their decision to join or not. This corresponds more closely to the actual process of treaty formation, where countries typically decide the commitments of the coalition members and what triggers entry into force *before* they decide whether to ratify (join).

[6] A plurality voting rule is often implemented in local and national elections to determine a single winner when there are more than two candidates. See Myerson and Weber (1993) for an analysis of voting equilibria under plurality voting rules.

In stage two of the coalition formation game (*the coalition stage*), the players decide independently and sequentially whether or not to join the coalition.[7] The order in which the players make their join/not join decisions is unknown in the voting stage of the game, and each potential order is equally likely. This order is revealed to the players in the coalition stage. Decisions to join or not join are observed by all other players in the second stage.[8] If enough players join so that the membership requirement is met, the coalition forms and its members provide their units of the public good. Throughout we call such a coalition an *effective coalition*. Those that do not join an effective coalition do not provide their units of the public good (because $c > b$), but still benefit from its provision. If the minimum membership requirement is not met, an effective coalition does not form and no player contributes to the public good.

We look for a subgame perfect equilibrium for this game, consisting of individual votes for the minimum membership requirement in the first stage of the game and decisions to join or not join a coalition in the second stage. An equilibrium is found by backward induction, so we start by describing the coalition stage. At this point in the game a minimum membership requirement has been chosen. Denote the membership requirement as s_p. Using (1), those who decide to join a coalition with s members earn:

$$\pi^m(s) = \begin{cases} A + bs - c & \text{for } s \leq \bar{s} \text{ if } s \geq s_p; \\ A + b\bar{s} - c & \text{for } s > \bar{s} \text{ if } s \geq s_p; \\ A & \text{if } s < s_p. \end{cases} \tag{2}$$

Throughout the superscript m indicates the player is a member of a coalition. Nonmembers, identified by the superscript nm, earn:

$$\pi^{nm}(s) = \begin{cases} A + bs & \text{for } s \leq \bar{s} \text{ if } s \geq s_p; \\ A + b\bar{s} & \text{for } s > \bar{s} \text{ if } s \geq s_p; \\ A & \text{if } s < s_p. \end{cases} \tag{3}$$

From (2) and (3) it is clear that when an effective coalition forms, nonmembers earn strictly higher profits than members, because they enjoy the benefits of public good contributions without incurring the cost of contributing. However, if the minimum membership requirement is not met, both members and nonmembers earn their noncooperative payoff A.

Players will join an effective coalition only if it is profitable for them to do so in the sense that they earn at least as much in the coalition as they would if no coalition formed. Thus, a coalition is profitable for its members if and only if $\pi^m(s|s \geq s_p) = A + bs - c \geq A$. Assume throughout that a coalition with $\bar{s}$ members is strictly profitable so that $\pi^m(\bar{s}|s \geq s_p) > A$, or rather, $b\bar{s} - c > 0$. Note from (2) that $\pi(s|s \geq s_p)$ is increasing in s up to $\bar{s}$ and then is constant. Combined with $b < c$, these relations imply that there exist coalition sizes that are strictly greater than one and weakly less than $\bar{s}$ that are profitable. The smallest of these profitable coalitions is:

$$s_{\min} = \min\left\{s | \pi^m(s | s \geq s_p) = A\right\} = \min\left\{s | s \geq c/b\right\}, \tag{4}$$

---

[7] Although our game is motivated by the model of Carraro et al. (2009), it is not a special case of their model. This is primarily due to the facts that they require unanimity in the vote for the membership requirement and players make their decision to join an agreement simultaneously.

[8] In fact, having the players decide to join or not in sequence and with perfect information about these decisions is a reasonable description of the actual process of treaty accessions. Differences in the lengths of national debates about the decision to ratify a treaty and differing positions of a treaty on national legislative agendas imply that ratification decisions must be sequential. Moreover, the decision to ratify a treaty or not tends to be very public.

where $s_{\min} \in (1, \bar{s}]$. For an effective coalition to form in the second stage of the game both the membership requirement and the profitability requirement must be satisfied.

The players make their decisions to join or not join sequentially. Define a *critical* player as one whose choice to not join a coalition prevents the coalition from forming. That is, a player is critical if and only if $n - s_p$ players have already opted out of the coalition. Suppose at first that the membership requirement is $s_p \geq s_{\min}$. A critical player in this situation will always choose to join a coalition of s_p members. To see why, note first that if a critical player refuses to join a coalition then no coalition will form and the player will earn A. However, the player earns a higher payoff if he is part of a profitable coalition. So, if a critical player joins a coalition then he either earns this higher payoff if all the other critical players also join, or he earns the payoff A if one of the other critical players refuses to join and a coalition does not form. Therefore, when $s_p \geq s_{\min}$ a critical player will always join a coalition because he cannot be worse off by doing so and he may be strictly better off. Since all critical players will join a coalition, a noncritical player would never join, because it is always more profitable to stay out of a profitable coalition. Therefore, given $s_p \geq s_{\min}$, the outcome of the second stage coalition game is that an effective coalition of s_p individuals will form. Moreover, since nonmembers earn strictly higher payoffs than members, the first $n - s_p$ players in the sequence will decide to not join the coalition and the last s_p players will decide to join. In the second stage of the game, all members of the coalition make their contributions to the public good, while the remaining players contribute nothing.

Now suppose that the membership requirement is $s_p < s_{\min}$. In this case, a critical player will choose to not join because he would be joining an unprofitable coalition. Since no critical player will join a coalition if $s_p < s_{\min}$, the outcome of the second stage coalition game in this case is that an effective coalition will not form and all players contribute nothing to the public good.

To sum up, the equilibrium of the second stage coalition formation game is a coalition size and provision of the public good s^* such that $s^* = s_p$ if $s_p \geq s_{\min}$ and $s^* = 0$ if $s_p < s_{\min}$. In the first case the first $n - s^*$ players to decide whether or not to join the coalition choose not to, while the last s^* players join the coalition.[9]

Now let us examine the determination of the minimum membership requirement s_p in the voting (first) stage of the game. Following Carraro et al. (2009), we assume that players vote sincerely in the sense that they vote for the membership requirement that gives them the highest expected payoff from the perspective of the first stage of the game.[10]

From our description of the coalition stage of the game, if the membership requirement is not less than the minimum profitable coalition, a player's payoff is determined by his position in the order in which the players make their decisions whether to join a coalition. Recall that this sequence is unknown in the first stage of the game. Hence, a player that expects an effective coalition of s players to form in the second stage evaluates the probability that he will be a member of this coalition as s/n and the probability that he will not be a member as $(n - s)/n$. His expected payoff from the perspective of the first stage of the game is then:

---

[9] Others have examined sequential decision making in threshold public good games. Erev and Rapoport (1990) and Cooper and Stockman (2002) derive similar equilibria to ours in threshold public good games in which players are assigned the order in which they must decide whether to contribute to a public good. In their studies a threshold is specified exogenously, which is referred to as the minimum contributing set. We noted the similarity between minimum contributing sets and minimum membership requirements for international treaties in the introduction. McEvoy (2010) explores the endogenous order of sequential decisions in public good games and finds that the timing of participation decisions is sensitive to the threshold in these games. In particular, he finds that subjects are more likely to rush to opt out of voluntary coalitions when the free-riding payoff is larger.

[10] Palfrey (1984) and Osborne and Slivinski (1996) also analyze games given sincere voting.

$$v(s) = (s/n)\,\pi^m(s) + ((n-s)/n)\,\pi^{nm}(s). \tag{5}$$

Remove the s_p constraint from (2) and (3), substitute the results into (5), and differentiate with respect to s to obtain:

$$v'(s) = \begin{cases} b - c/n > 0, & \text{for } s \le \bar{s}; \\ -c/n < 0, & \text{for } s > \bar{s}. \end{cases} \tag{6}$$

Equation (6) indicates that the expected payoff of every player is maximized at the efficient coalition size. Given that all players vote sincerely in the first stage of the game, each of them votes to implement the efficient coalition as the minimum membership requirement. That is, $s_p = \bar{s}$.

Moving ahead to the coalition formation stage, since the efficient coalition size is profitable, $s_p = \bar{s} \ge s_{\min}$. Therefore, the coalition $s^* = \bar{s}$ will form. Obviously, if $\bar{s} = n$, then all players will join the coalition. If $\bar{s} < n$, then the first $n - \bar{s}$ players to make their join/not join decision will choose to not join, while the last $\bar{s}$ players join the coalition.[11] The main prediction of the theory is that the endogenous determination of a minimum membership requirement allows players to form an efficient coalition to provide a public good. In the next section we present our experimental design used to test this prediction under two scenarios; one in which efficiency requires full participation (i.e., the grand coalition) and one in which efficiency requires only partial participation.

3 Experiments

Following the theoretical model, our experiments have two stages, a voting stage and a coalition formation stage. The context of the experiment is fairly generic; subjects decide whether to join a coalition (called an agreement) with its members making a discrete choice to contribute to the public good (called a public account). We implement two experimental treatments that differ according to the returns to individual contributions to the public good. In one treatment the marginal return to contributions is constant so that the efficient coalition is the grand coalition. In the other treatment, the marginal return to public good contributions is constant up to an aggregate level and then zero so that the efficient coalition is smaller than the grand coalition.

We chose parameter values of $n = 6$, $A = 10$, and $c = 10$ for each treatment. In one treatment the marginal return to all individual contributions was set at $b = 4.5$. Since $nb - c > 0$ for all contributions, the efficient coalition in this case is six individuals ($\bar{s} = 6$). In the other treatment, $b = 4.5$ for contributions up to three units and then zero after three units. The efficient coalition size in this case is three individuals ($\bar{s} = 3$). Plugging our parameter values into Eq. (4) reveals a minimum profitable coalition size of three members. Therefore, in both treatments, coalitions of three or more are profitable for the members.

All sessions were run at the experimental economics laboratory at Appalachian State University using software specifically designed for this experiment.[12] Subjects were recruited from the undergraduate and graduate student populations. Two sessions were implemented for each of the treatments. In each session, three groups of six subjects were in the lab.

---

[11] Carraro et al. (2009) provide sufficient conditions for their game to result in the formation of the grand coalition. Our game satisfies those sufficient conditions. Carraro et al. do not examine the case in which the efficient coalition is smaller than the grand coalition.

[12] The experiment instructions can be found at: http://davemcevoy.weebly.com/uploads/2/2/7/0/2270780/instructions_ere2014.pdf.

These groups of six were reshuffled each period throughout the experiment, which lasted 20 periods.[13] For each treatment we have 720 individual-level observations, which include their votes for the membership requirement in the first stage and their decisions to join a coalition in the second stage. We also have 120 group-level observations on coalition formation, public good provision and overall group efficiency. The sessions lasted for roughly 1 h, subjects were paid their cumulative earnings and earned $16.50 on average (20 experimental dollars = $1).

In the first stage of the experiments, subjects simultaneously voted on the minimum membership requirement for a coalition to form in the second stage. The membership requirement that received the most votes was implemented in the second stage of the game. Ties were settled by a random draw.

In the coalition formation stage, each subject decided whether or not to join the coalition. Following the theoretical model, these decisions were made sequentially under perfect information. Each of the six subjects in a group was randomly assigned an order in which to make decisions, and this order changed each period.[14] For example, if a subject was assigned an order number of three, then she had to wait for the first two players to submit their choices before she could make her decision. Once a subject made their choice they could not change it. During this stage, all subjects were informed about whether the subjects before them joined the agreement or did not join the agreement. They were also constantly reminded of the membership requirement for the agreement to form. Therefore, each player knew with certainty whether their participation decision was critical for the coalition to form.

If enough players joined the coalition to satisfy the minimum membership requirement, then the coalition 'formed' and those that joined contributed to the public good. Those who did not join did not contribute. If too few subjects joined the coalition to satisfy the membership requirement, then no coalition formed and no one contributed to the public good.

4 Results

Our experimental data suggest the following broad conclusions. The majority of subjects in both treatments voted for efficient minimum membership requirements. These votes led to the adoption of the efficient membership requirements in roughly three out of every four trials for both treatments. When efficiency required full participation (the $\bar{s} = 6$ treatment), coalitions formed in 85 % of all trials. The efficient coalition almost always formed (91.3 %) when the efficient membership requirement was adopted. Overall, efficiency was very high in this treatment, about 87.4 % as measured by the ratio of realized group earnings to maximum group earnings. Hence, group results in the $\bar{s} = 6$ treatment closely matched our theoretical predictions that groups will use the membership requirement to form efficient coalitions. However, performance was significantly worse in the treatment for which a 3-player coalition was efficient (the $\bar{s} = 3$ treatment). In this treatment coalitions formed in only 61 % of all trials and the efficient coalition formed in only about half the trials. Consequently, average efficiency was significantly lower in these treatments (about 80 %) than in the $\bar{s} = 6$ treatments. Interestingly, the efficient coalition was blocked about a third of the time when the efficient membership requirement was adopted. This behavior is not
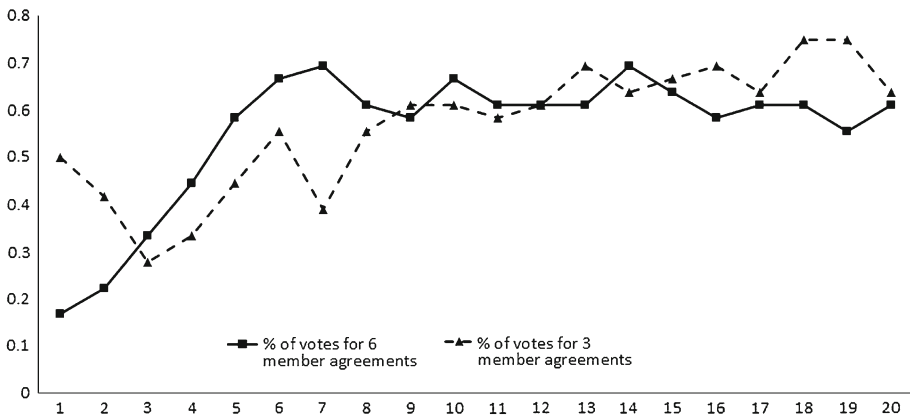
Table 1 Individual votes and referenda outcomes by minimum membership requirement and treatment

| | Minimum membership requirement | | | | | | |
	1	2	3	4	5	6	Totals
$\bar{s} = 6$	35 (4.9%)	27 (3.8%)	53 (7.4%)	66 (9.2%)	139 (19.3)	**400 (55.6%)**	720
	2 (1.7%)	3 (2.5%)	6 (5.0%)	7 (5.8%)	10 (8.3%)	**92 (76.7%)**	120
$\bar{s} = 3$	91 (12.6%)	153 (21.2%)	**409 (56.8%)**	32 (4.4%)	9 (1.3%)	26 (3.6%)	720
	7 (5.8%)	19 (15.8%)	**92 (76.7%)**	1 (0.8%)	0 (0.0%)	1 (0.8%)	120

Top of each cell Number of votes for each minimum membership requirement (percent of total votes by treatment). Note there are 720 individual votes per treatment. *Bottom of each cell* Number of times each minimum membership requirement was implemented (percent of total trials by treatment). Note there are 120 group-level observations per treatment

Fig. 1 Percentage of votes for efficient coalitions by treatment

consistent with a theory of coalition formation that includes only individuals with standard payoff-maximizing preferences. However, we will demonstrate that this behavior is consistent with a theory that allows for preferences over inequality.

We begin our detailed analysis of the results by examining the data on voting for the membership requirement in the first stage of the experiments. Table 1 provides votes and referenda outcomes by membership requirement and treatment. The first row in each cell contains the number of votes and percentage of total votes (out of 720 for each treatment) for that minimum membership requirement. The second row in each cell contains the number of times and percentage of trails in the treatment (out of 120) that membership requirement was implemented.

Under the $\bar{s} = 6$ treatments, the 6-player membership requirement received 55.6% (400 of 720) of total votes, which is considerably more than the 19.3% received by the second most preferred option of a 5-player membership requirement. The remaining four options received even fewer votes. This voting behavior resulted in the selection of the efficient membership requirement in 76.7% (92 of 120) of referenda.

The percentage of votes for efficient membership requirements is very similar across the two treatments. In the $\bar{s} = 3$ treatment, the 3-player membership requirement received 56.8% of the votes. Membership requirements of size two received the second most with 21.2% of votes. Mirroring the results from the $\bar{s} = 6$ treatment, voting led to the adoption of the

Table 2 Regression results comparing the two treatments

	Vote on efficient MMR	Efficient referenda outcomes	Total coalition formation	Efficient coalition formation	Efficiency
$\bar{s} = 3$	0.029	0.006	−0.264	−0.268	−0.069
	0.817	0.922	0.000	0.003	0.015
Model	131.69	24.83	44.27	22.97	3.20
	0.000	0.073	0.001	0.115	0.000
Period effects	109.54	18.57	20.56	5.89	2.58
(Chi-square)	0.000	0.182	0.2468	0.970	0.001
Subject effects	665.64	–	–	–	–
(Chi-square)	0.000				
n	1440	240	240	184	240

Notes: The first four columns report average marginal effects (and p values) for the $\bar{s} = 3$ treatment (relative to the $\bar{s} = 6$ treatment) from Probit regressions. The first model controls for period fixed effects and subject random effects. The other three Probit regressions are at the group-level and control for period fixed effects. The last model is a linear regression with a dependent variable defined as total group earnings/max group earnings and controls for period fixed effects

efficient membership requirement in 76.7 % (92 of 120) of referenda. Figure 1 illustrates the percentage of votes for efficient membership requirements over periods for both treatments. To summarize, the majority of subjects in both treatments voted for efficient membership requirements, and these were adopted in roughly three out of every four trials.

Figure 1 suggests that the time series of the percentages of votes for the efficient coalition are very similar between the two treatments. We confirm this with results from a Probit model conditioning individual voting decisions on treatment, period fixed effects and subject-specific random effects. These results are reported in the first column of Table 2. Note that the treatment effect ($\bar{s} = 3$) is insignificant. In addition, the Probit model in the second column of Table 2 confirms that the likelihood of groups adopting the efficient membership requirement is statistically equivalent in the two treatments.

Recall that our theoretical model yields the prediction that expected-payoff-maximizing players would vote for minimum membership requirements equal to efficient coalition sizes. Our experimental results from the voting stage are broadly consistent with this prediction in that the majority of subjects voted to adopt the efficient membership requirements. However, in contrast with our theoretical model we do observe significant, though minority, voting for smaller-than-efficient membership requirements. Our theoretical model is not helpful in explaining why 45 % of votes were cast for inefficient agreement sizes. It is interesting to note that in each treatment the second most frequent vote was for one member less than the efficient agreement size. These votes are consistent with players attempting to increase their chances to free ride on profitable agreements. For example, a lone free rider on a coalition of five players in the full participation treatment would be the highest possible earner. Of course there are other possible explanations for this behavior. If players are not confident that other group members will join the agreement when it is rational to do so (if for example, they have a trembling hand when making their participation decision), then a lower participation threshold may appear less risky.[15]

---

[15] We thank an anonymous referee for suggesting this possible explanation. See McGinty (2011) for an example of trembling hand equilibria in a coalition formation game.

Table 3 Coalition formation by minimum membership requirement and treatment

	Minimum membership requirement						
	1	2	3	4	5	6	Total
$\bar{s} = 6$	1	1	5	4	7	**84**	102
	0.8%	0.8%	4.2%	3.3%	5.8%	**70.0%**	85.0%
	50.0%	33.3%	83.3%	57.1%	70.0%	**91.3%**	
$\bar{s} = 3$	1	9	**62**	0	0	1	73
	0.8%	7.5%	**51.7%**	0.0%	0.0%	0.8%	60.8%
	16.7	47.7%	**67.4%**	0.0%	0.0%	100.0%	

Top of each cell Number of times coalitions formed. *Middle of each cell* Percentage coalition formation by number of trials per treatment. *Bottom of each cell* Percentage coalition formation by adopted membership requirement

While the results from the voting stage are equivalent in the two treatments, the results from the coalition formation stage differ significantly. Table 3 contains results concerning coalition formation by minimum membership requirement and treatment. For each membership requirement/treatment combination we provide the number of times a coalition formed under the membership requirement, this number as a percentage of total trials, and coalition formations as a percentage of times the membership requirement was adopted. The final column in Table 3 contains the number and percentage of trials a coalition of any size formed (i.e., all effective coalitions).

Under the $\bar{s} = 6$ treatment, coalitions of any size formed in 102 of 120 trials (85%). In 84 of 120 trials the efficient coalition formed (70%). Moreover, when players adopted the efficient membership requirement of six players, coalitions almost always formed (84 of 92, 91.3%). Thus, coalitions formed quite frequently in the $\bar{s} = 6$ treatments and most coalitions were efficient. Other coalition sizes formed much less frequently. This suggests that the main reason why coalitions did not form in a subset of trials is because groups sometimes failed to implement the 6-player membership requirement. In both treatments, if a coalition failed to form then no player contributed to the public good and each earned 10 experimental dollars.

In contrast, coalitions formed significantly less frequently under the $\bar{s} = 3$ treatment. Overall, coalitions formed in 60.8% of trials in this treatment, which is significantly lower than the 85% coalition formation rate under the $\bar{s} = 6$ treatment. This is confirmed by the results of the third model in Table 2, which is a Probit model of total coalition formation. Note that the dummy for the $\bar{s} = 3$ treatment is negative and significant. Furthermore, the efficient coalition formed in 51.7% of trials (62 of 120), compared with 70% of trials in the $\bar{s} = 6$ treatment. The most striking difference between the two treatments is the percentage of coalitions that formed conditional on the adoption of an efficient membership requirement. In both treatments, the referenda resulted in efficient membership requirements in 92 of 120 trials. With the $\bar{s} = 6$ treatment, coalitions formed in 84 of those 92 trials (91.3%). However, under the $\bar{s} = 3$ treatment, coalitions formed in 62 of those 92 trials (67.4%). The fourth Probit model in Table 2 confirms that the likelihood of a coalition forming when the efficient membership requirement was adopted is significantly lower in the $\bar{s} = 3$ treatment.

Let us explore this phenomenon more closely, since it implies that some individuals willfully blocked efficient coalitions in the sense that they were critical for the formation of the coalition but refused to join so the coalition failed to form. The 3-player membership requirement was implemented in 92 out of 120 trials (76.7% from Table 1) under the $\bar{s} = 3$

Table 4 Public good provision and efficiency	Treatment	Average public good provision	Efficiency
	$\bar{s} = 6$	4.75	87.4 %
		(0.26)	(2.21)
Standard errors are in parentheses. Each treatment consists of 120 group-level observations	$\bar{s} = 3$	1.78	79.6 %
		(0.31)	(2.07)

treatment. The efficient coalition failed to form in 30 of these trials. These blocks are inconsistent with a model that contains only individuals with standard expected-payoff-maximizing preferences. Such individuals would never block a profitable coalition, and efficient coalitions are always profitable. However, these blocks are consistent with the presence of subjects who are averse to disadvantageous inequality. When $\bar{s} = 3$, the efficient outcome requires that the three non-members earn strictly higher payoffs than the three members. If players dislike payoff inequality, it is possible that coalitions that are profitable in terms of material payoffs are no longer profitable in terms of utility. In fact, minimum profitable coalition sizes for inequality-averse players will be weakly greater than those for players with standard preferences. Therefore, inequality-averse players may block efficient coalitions that are not individually profitable given their inequality preferences. That inequality averse individuals require higher minimum profitable coalition sizes is demonstrated theoretically in Kosfeld et al. (2009) for a game that is similar to ours. We demonstrate this result for our model in the Appendix.

We complete our data analysis with results on average public good provision and average efficiency in Table 4. Efficiency for each group in each period is calculated as the ratio of aggregate payoffs to maximum attainable payoffs. As expected, public good provision and efficiency were lower in the $\bar{s} = 3$ treatment than in the $\bar{s} = 6$ treatment. The significance of the difference in efficiency levels is confirmed in the linear regression results contained in the final column of Table 2.

To judge the performance of the endogenous determination of minimum membership requirements, we can compare the efficiency level of our $\bar{s} = 6$ treatment (87.4 %) to recent experimental coalition formation studies that do not include this feature. The efficiency level in this treatment is quite high relative to these other studies.[16] For example, Kosfeld et al. (2009) conducted two coalition formation treatments which differed in the minimum number of members required for coalitions to be profitable (either 3 or 2 out of 4). As noted earlier, subjects in their experiments first decided whether to join a coalition, and then in a second stage the coalition members voted whether to contribute all of their endowment to the public good or not. Efficiency levels in their experiments were 51 and 70 % for minimum profitable coalition sizes of 3 and 2, respectively.

Among a number of treatments in McEvoy et al. (2011) were two coalition formation treatments that differed in terms of the individual costs of providing the good, producing different sizes of minimally profitable coalitions. These were then set as minimum membership requirements. One treatment required 3 of 10 subjects to join before a coalition formed while the other had a membership requirement of 6 of 10 subjects. The authors did not allow contributions from nonmembers and did not allow contributions if a coalition did not form; hence their efficiency measures are fairly conservative. Efficiency levels were 59.6 % for the 3-player membership requirement and 56.2 % for the 6-player membership requirement.

---

[16] We cannot use the literature to judge the performance of our treatment, because all other coalition formation experiments require the formation of the grand coalition for efficient provision of a public good.

Subjects performed much worse in the experiments of Dannenberg et al. (2010). They conducted three coalition formation treatments with 10-player groups. The authors report an efficiency level of 24 % for one treatment in which coalition members' contribution decisions were determined automatically by maximizing joint welfare. In another treatment coalition members only contributed half of what was required to maximize joint welfare and efficiency fell to 12 %. The efficiency level was 29 % in a third treatment in which coalition members endogenously determined the required minimum contribution. Although there are significant differences in the protocols between the experiments in Dannenberg et al. (2010) and our own (including group sizes), the comparison suggests that endogenous thresholds serve as effective coordination devices in coalition formation games.

5 Conclusion

Many international environmental agreements include a minimum membership requirement for entry into force. Despite the wide-spread use of these requirements, little is known about their effectiveness. We have analyzed a coalition formation game that includes the endogenous determination of a minimum membership requirement and then tested the theoretical predictions of the game using a series of laboratory experiments. The main prediction of our theoretical model is that agents will vote to adopt efficient coalition sizes as membership requirements and that these coalitions will form. Our experimental results are largely consistent with the theoretical predictions when efficiency requires the grand coalition to form. In this treatment, coalitions formed 85 % of the time and over 90 % of these were the efficient grand coalition. These results demonstrate the value of the endogenously determined membership requirements. In fact, the level of efficiency for this treatment is significantly higher than in other coalition formation experiments that do not include the endogenous determination of membership requirements.

However, in our treatment for which the efficient coalition required only a subset of the group, coalitions formed only 60 % of the time and efficient coalitions were deliberately blocked about a third of the time. If these blocks had not occurred, performance in this treatment would have matched the performance of the treatment that required the grand coalition to form for efficient provision of the public good. Although individuals with standard preferences would never block efficient coalitions from forming, these actions are consistent with inequality-averse individuals. Equity concerns may help explain the choice of membership requirements in many existing voluntary institutions; in particular, the fact that many international agreements require very high levels of participation (Barrett 2003). In light of our results it is possible that equity concerns can limit the formation of international agreements when membership requirements would allow free riders. One might expect that inefficiently high membership requirements would emerge to limit freeriding, although we do not observe this in our experiments. The role of equity preferences in the formation of voluntary agreements is not well understood, so further research in this area would be beneficial. Another caveat is that our analysis assumes homogeneity in the costs and benefits of providing a public good. Future research should relax this simplifying assumption and explore the role of endogenous minimum participation requirements among heterogeneous players.

Our objective was to empirically test a feature of international governance—endogenous minimum participation—that is included in almost all environmental treaties. Although our laboratory experiments necessarily abstract from the many intricacies involved with international management of shared resources, the results shed light on the effectiveness of the institution in a controlled environment. The take away message for stakeholders in interna-

tional diplomacy is that endogenous minimum participation requirements can be effective at fostering cooperation, but may be significantly less effective when targeted participation allows for free riding.

Appendix

Here we incorporate inequality aversion into our model to demonstrate that minimum profitable coalitions are weakly larger for inequality averse individuals than for individuals with standard preferences. We follow Fehr and Schmidt (1999) in modeling preferences over inequality. Suppose at first that $\bar{s} = n$ and $s_p = 0$. Given the financial payoffs (2) and (3) with these restrictions, define the utility of a member of an effective coalition with s members as:

$$u_i^m(s) = \pi_i^m(s) - \frac{\alpha_i}{n-1} \sum_{j \neq i} \max\left(\pi_j - \pi_i^m(s), 0\right) - \frac{\beta_i}{n-1} \sum_{j \neq i} \max\left(\pi_i^m(s) - \pi_j, 0\right), \quad (7)$$

where $\alpha_i > 0$ captures the player's loss from disadvantageous inequality and $\beta_i > 0$ captures her loss from advantageous inequality. Since $\pi_j^{nm}(s) - \pi_i^m(s) = c$ and $\pi_j^m(s) - \pi_i^m(s) = 0$ from (2) and (3), (7) can be written as:

$$u_i^m(s) = A + bs - c - \frac{\alpha_i c(n-s)}{n-1}. \quad (8)$$

Similarly, the utility of a nonmember of an effective coalition with s members is:

$$u_i^{nm}(s) = \pi_i^{nm}(s) - \frac{\alpha_i}{n-1} \sum_{j \neq i} \max\left(\pi_j - \pi_i^{nm}(s), 0\right)$$

$$- \frac{\beta_i}{n-1} \sum_{j \neq i} \max\left(\pi_i^{nm}(s) - \pi_j, 0\right), \quad (9)$$

which can be written as

$$u_i^{nm}(s) = A + bs - \frac{\beta_i sc}{n-1}. \quad (10)$$

It is straightforward to show that the free-riding incentive is preserved in this model if $(n-1)(\beta_i - 1) < \alpha_i$; that is, as long as the aversion to advantageous inequality is not too strong relative to the aversion to disadvantageous inequality. Incorporating the efficient coalition size $\bar{s} \leq n$ to determine individual payoffs yields:

$$u_i^m(s) = \begin{cases} A + bs - c - \frac{\alpha_i(n-s)c}{n-1} & \text{for } s \leq \bar{s}; \\ A + b\bar{s} - c - \frac{\alpha_i(n-s)c}{n-1} & \text{for } s > \bar{s}; \end{cases} \quad (11)$$

$$u_i^{nm}(s) = \begin{cases} A + bs - \frac{\beta_i sc}{n-1} & \text{for } s \leq \bar{s}; \\ A + b\bar{s} - \frac{\beta_i sc}{n-1} & \text{for } s > \bar{s}. \end{cases} \quad (12)$$

Using (11), an individual's minimum profitable coalition size can be characterized as:

$$\tilde{s}_i \begin{cases} = \hat{s}_i \text{ if } \hat{s}_i \leq \bar{s} \\ > \hat{s}_i \text{ if } \hat{s}_i > \bar{s}, \end{cases} \quad \text{where} \quad \hat{s}_i = \frac{c(n-1) + \alpha_i cn}{b(n-1) + \alpha_i c}. \quad (13)$$

To demonstrate $\tilde{s}_i$, we first derive $\hat{s}_i$ as the solution to:

$$A + bs - c - \frac{\alpha_i (n - s)c}{n - 1} = A. \qquad (14)$$

Since $u_i^m(s)$ in (11) is increasing in s, if $\hat{s}_i \leq \bar{s}$ then $\hat{s}_i$ is i's minimum profitable coalition size. However, if $\hat{s}_i > \bar{s}$, then $\tilde{s}_i$ must be the solution to

$$A + b\bar{s} - c - \frac{\alpha_i (n - s)c}{n - 1} = A. \qquad (15)$$

Plug $\hat{s}_i$ into (14) and $\tilde{s}_i$ into (15), set the resulting equations equal to each other and collect terms to obtain $b(\hat{s}_i - \bar{s}) = \alpha_i c(\tilde{s}_i - \hat{s}_i)/(n - 1)$, which implies that $\tilde{s}_i > \hat{s}_i$ if $\hat{s}_i > \bar{s}$.

Recall from (4) that the minimum profitable coalition size for an individual with standard preferences is $s_{\min} = \min\{s | s \geq c/b\}$. From (13), $\tilde{s}_i \geq \hat{s}_i$, $\hat{s}_i$ is increasing in α_i, and $\hat{s}_i = c/b$ for $\alpha_i = 0$. Together, these imply $\tilde{s}_i \geq s_{\min}$ for an individual with disadvantageous inequality aversion (i.e., $\alpha_i > 0$). Therefore, such an individual has a weakly higher minimum profitable coalition size than an individual with standard preferences.

By substituting in our experimental parameters into equation (13) we can demonstrate how large α_i must be to increase the minimum profitable coalition size beyond the efficient size. In the treatment with an efficient coalition of three members, if α_i exceeds 0.584, then the minimum profitable coalition for an individual will exceed three members. As a frame of reference, at least 40 % of players in Fehr and Schmidt's analysis were estimated to have $\alpha > 0.50$.

References

Barrett S (1994) Self-enforcing international environmental agreements. Oxf Econ Pap 46(5):878–894

Barrett S (2003) Environment and statecraft: the strategy of environmental treaty making. Oxford University Press, Oxford

Barrett S, Dannenberg A (2012) Climate negotiations under scientific uncertainty. Proc Natl Acad Sci 109(43):17372–17376

Black J, Levi MD, de Meza D (1993) Creating a good atmosphere: minimum participation for tackling the 'Greenhouse Effect'. Economica 60(239):281–293

Cadsby CB, Maynes E (1999) Voluntary provision of threshold public goods with continuous contributions: experimental evidence. J Pub Econ 71(1):53–73

Carraro C, Siniscalco D (1992) The international dimension of environmental policy. Eur Econ Rev 36(2–3):379–387

Carraro C, Marchiori C, Oreffice S (2009) Endogenous minimum participation in international environmental treaties. Environ Resour Econ 42(3):411–425

Cason TN, Plott CR (1996) EPA's new emissions trading mechanism: a laboratory evaluation. J Environ Econ Manag 30(2):133–160

Cooper DJ, Stockman CK (2002) Learning to punish: experimental evidence from a sequential step-level public goods game. Exp Econ 5(1):39–51

Dannenberg A, Lange A, Sturm B (2010) On the formation of coalitions to provide public goods—experimental evidence from the lab. NBER working paper no. 15967

Erev I, Rapoport A (1990) Provision of step-level public goods: the sequential contribution mechanism. J Confl Resolut 34(3):401–425

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. Q J Econ 114(3):817–868

Fehr E, Gächter S (2000) Cooperation and punishment in public goods experiments. Am Econ Rev 90(4):980–994

Gurerk O, Irlenbusch B, Rockenbach B (2006) The competitive advantage of sanctioning institutions. Science 312(5770):108–11

Harstad B (2006) Flexible integration? Mandatory and minimum participation rules. Scand J Econ 108(4):683–702

Kosfeld M, Okada A, Riedl A (2009) Institution formation in public goods games. Am Econ Rev 99(4):1335–1355

Kroll S, Cherry TL, Shogren JF (2007) Voting, punishment, and public goods. Econ Inquiry 45(3):557–570

Lange A, Vogt C (2003) Cooperation in international environmental negotiations due to a preference for equity. J Pub Econ 87(9–10):2049–2067

McEvoy D (2010) Not it: opting out of voluntary coalitions that provide a public good. Pub Choice 142(1):9–23

McEvoy D, Murphy J, Spraggon J, Stranlund J (2011) The problem of maintaining compliance with stable coalitions: experimental evidence. Oxf Econ Pap 63(3):475–498

McGinty M (2011) A risk-dominant allocation: maximizing coalition stability. J Pub Econ Theory 13(2):311–325

Myerson RB, Weber RJ (1993) A theory of voting equilibria. Am Polit Sci Rev 87(1):102–114

Osborne M, Slivinski A (1996) A model of political competition with citizen-candidates. Q J Econ 111(1):65–96

Palfrey T (1984) Spatial equilibrium with entry. Rev Econ Stud 51(1):139–156

Plott C (1987) Dimensions of parallelism: some policy applications of experimental methods. In: Roth A (ed) Laboratory experimentation in economics: six points of view. Cambridge University Press, Cambridge

Rapoport A (1985) Provision of public goods and the MCS experimental paradigm. Am Polit Sci Rev 79(1):148–155

Ringius L, Torvanger A, Underdal A (2002) Burden sharing and fairness principles in international climate policy. Int Environ Agreem Polit Law Econ 2(1):1–22

Shogren JF, Hurley TM (1999) Experiments in environmental economics. In: van den Bergh JCJM (ed) Handbook of environmental and resource economics. Edward Elgar, Cheltenham

Stranlund JK, Murphy JJ, Spraggon JM (2011) An experimental analysis of compliance in dynamic emissions markets. J Environ Econ Manag 62(3):414–429

Sutter M, Haigner S, Kocher MG (2010) Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. Rev Econ Stud 77(4):1540–1566

Tyran J-R, Feld LP (2006) Achieving compliance when legal sanctions are non-deterrent. Scand J Econ 108(1):135–156

Walker JM, Gardner R, Herr A, Ostrom E (2000) Collective choice in the commons: experimental results on proposed allocation rules and votes. Econ J 110(460):212–234

Weikard H-P, Wangler L, Freytag A (2009) Minimum participation rules with heterogeneous countries. Jena Economic Research Papers 2009–077, ISSN 1864–7057