Not It: Opting out of Voluntary Coalitions that Provide a Public Good

DAVID M. MCEVOY

Appalachian State University

Economics Department

Raley Hall, 416 Howard Street

ASU Box 32051

Boone, NC 28608

Phone: 828-262-6126

Fax: 828-262-6105

Email:mcevoydm@appstate.edu

Acknowledgements: The author would like to thank John Stranlund, John Spraggon, James Murphy, Mike McKee, Todd Cherry and various attendees of the 2008 international meeting of the Economic Science Association for their helpful comments on this research.

Abstract: Most coalitions that form to increase contributions to a public good do not require full participation by all users of the public good, and therefore create incentives for free riding. If given the opportunity to opt out of a voluntary coalition, in theory, agents should try to be among the first to do so, forcing the remaining undecided agents to bear the cost of participating in the coalition. This study tests the predicted sequence of participation decisions in voluntary coalitions using real-time threshold public goods experiments. We find that subjects' behavior is more consistent with the theoretical predictions when the difference in payoffs between coalition members and free-riding non-members is relatively large.

JEL Codes: *H41, C92, C72, Q50*

Keywords: *voluntary coalitions, voluntary agreements, public goods experiments, free riding*

1. Introduction

It is well known that privately funded public goods will, in many cases, be underprovided because self-interested agents will attempt to benefit from public goods without bearing the cost of providing them. As a result, the provision of public goods constitutes a social dilemma in which individually rational choices are not collectively rational. One way to potentially increase provision of a public good is with a voluntary coalition, through which a group of agents jointly commit to increasing contributions to a public good. Such coalitions are appealing because in many situations agents would be willing to increase their contributions to a public good provided that others (or a subset of others) commit to doing the same. Coalitions of this nature can describe countries within international agreements to protect the environment, individuals when making charitable contributions, firms involved in domestic voluntary agreements and users of some common pool resources.

Most voluntary coalitions that provide public goods do not require full participation by all agents involved before they are implemented. For example, the Kyoto Protocol, the international environmental agreement developed to manage global greenhouse gas emissions, required ratification from at least 55% of the parties to the Convention and those parties needed to account for at least 55% of the total 1990 greenhouse gas emissions (UNFCCC 1998). In general, participation thresholds established under voluntary coalitions do not require all users of a public good to become members, and as a result, there are incentives for agents to remain outside of a coalition in the hopes of benefiting from the public good without bearing the cost of providing it. While all agents may benefit from an effective voluntary coalition (throughout, an *effective* coalition is one that leads to an increase in the provision of a public good), the non-members benefit disproportionately as a result of avoiding the cost of providing the public good. Thus, voluntary coalitions are susceptible to free-rider incentives that could potentially undermine their ability to increase contributions to a public good.

Related to the issue of free riding, one of the unique features of many voluntary coalitions is that, in general, the decision to join is made sequentially. A fixed window of time is typically established during which agents can decide voluntarily whether to become a member to a coalition.[1] The sequential decision making process allows for a more calculated form of free riding compared to when participation decisions are made simultaneously because agents know with certainty whether their decision to join will be critical for meeting a minimum participation requirement. Moreover, if agents are given the chance to formally opt out of a coalition, in theory, rational agents should try to be among the first to do so (i.e., choose not to join the coalition) in order to place the cost of participating on the remaining undecided agents. Although the game-theoretic prediction is clear, attempting to be among the first to opt out of a voluntary coalition is not the only strategy available. For example, is possible that agents may delay making their decision whether to participate - often referred to as a hold-out strategy - in the hopes that enough of the other agents will join before them thereby satisfying the participation requirement (Swanson 1999, p. 139; Barrett 2003, p. 142).

In most situations agents can deliberately opt out of voluntary coalitions. For example, the United States opted out of the Kyoto Protocol when the Bush Administration released to the public that it had no intention of ratifying the Protocol in 2005. However, with many coalitions agents only opt out *informally* by not actively joining, and therefore data are not available to test the theoretical hypothesis that agents will try to be among the first to opt out of voluntary coalitions.[2] Examining free-riding behavior in this setting is useful for guiding the design of effective participation rules that are fundamental to voluntary coalitions. This study analyzes a set of economic experiments in which subjects sequentially make explicit decisions whether to become members of a coalition.

Real-time, sequential public goods experiments are used to investigate the extent to which subjects deliberately opt out of joining coalitions that disproportionately benefit the free-riding non-members. The experiments utilize a discrete choice, threshold public good framework similar to that of Van de Kragt et al. (1983), Dawes et al. (1986) and Marks and Croson (1998), with a real-time

---

[1] See Barrett 2003 (pp. 133 – 164) for a discussion of the participation process with international environmental agreements. For a broad discussion of the participation process with various forms of domestic voluntary environmental agreements in the United States, Europe and Japan see Morgenstern and Pizer (2007).

[2] One notable exception to this rule is the process for implementing new regulations for members of the World Health Organization (WHO). According to Article 22 of the WHO Constitution, all new regulations, by default, come into force for each member unless a member explicitly rejects the regulation.

decision making processes similar to Dorsey (1992), Kurzban et al. (2001), Goren, Kurzban and Rapoport (2003) and Goren, Rapoport and Kurzban (2004). In this study, subjects make a single decision whether or not to join a voluntary coalition. Players have a fixed amount of time to make this decision while constantly being updated regarding the decisions made by the other group members. Once all subjects have made their decision, if the participation threshold is satisfied then the members are committed to contributing to the public good. Otherwise, the members do not contribute to the public good, no level of the public good is provided and members do not incur any costs for joining the coalition.[3] Throughout, the participation threshold is set to the smallest coalition of members required to make contributing to the public good profitable for the members. In other words, the participation threshold ensures that the members are better off with a coalition than without one.

Although previous experimental research has investigated the order of decisions made in sequential-choice, threshold public goods experiments (Erev and Rapoport 1990; Cooper and Stockman 2002; Goren, Kurzban and Rapoport 2003; Coats and Neilson 2005), these studies were not specifically designed to investigate the real-time sequence in which subjects opt out of mutually beneficial coalitions. Erev and Rapoport (1990) test the theoretical predictions of a threshold public goods game in which the order of decisions is determined exogenously. Group members make their decision to contribute to a public good in a defined sequence knowing with certainty the decisions made by the player or players before them. Their design, however, does not capture the endogenous nature of the voluntary participation (or contribution) process found in most threshold public good environments outside of the laboratory. On the other hand, the study by Gorzen, Kurzban and Rapoport (2003), does feature an environment in which decisions are made sequentially and the order of those decisions is endogenous. However, their design differs from our design in two fundamental ways. First, the players are *asymmetric* regarding their initial endowments, whereas the players in our study are symmetric. Second, and most importantly, the subjects begin each round in their experiment at the default position of not contributing to the public good and the only decision they can make is to change their position to contribute.

[3] This feature is similar to a money-back guarantee included in many experimental studies (Dawes et al. 1986; Isaac et al. 1989; Rapoport and Eshed-Levy 1989; Bagnoli and McKee 1991; Marks

Therefore, it is not possible to investigate the sequence of opting out behavior from their results.[4]

Two treatments are considered in this study, one in which the cost of contributing to the public good is relatively low (three of the ten players must join to satisfy the participation threshold, referred to as the *low-cost treatment*) and one in which the cost of contributing to the public good is relatively high (six of the ten players must join to satisfy the participation threshold, referred to as the *high-cost treatment*). A number of interesting results follow from this research. First, the percentage of trials in which an effective voluntary coalition formed (i.e., the participation threshold was either met or surpassed) was significantly lower in the *high-cost treatment* in which the participation threshold and the relative benefits from free riding were large compared to the *low-cost treatment* (62% vs. 96%). It is important to note that if a voluntary coalition did not form in either treatment it was because too many subjects chose to opt out of a coalition that would be mutually beneficial but disproportionately benefit free-riding non-members.

Second, the data show a complete reversal of the overall participation strategy implemented by the group members between the two treatments. Although exact equilibrium play was rare, the behavior of subjects in the *high-cost treatment*, in which the participation threshold and the relative payoffs to free riding were high, more closely resembled the theoretical predictions when compared to the *low-cost treatment*. In the *high-cost treatment*, the majority of the first decisions were made by subjects deliberately opting out of voluntary coalitions, while the reverse was true with the *low-cost treatment*. In summary, the sequence of subjects' participation decisions best matched the theoretical predictions when the difference in payoffs between coalition members and free-riding non-members was relatively large.

---

and Croson 1998; Rondeau et al. 1999; Cadsby et al. 2008).

[4] There are additional differences between our experiment and Gorzen, Kurzban and Rapoport (2003). Their study uses smaller group sizes (5 vs. 10 players) and the threshold in their experiment remains constant throughout while this study varies the threshold. Moreover, contributions in excess of the threshold were not utilized in their experiment whereas they are in this experiment. Finally, if the threshold was not satisfied in their experiment then players that contributed to the public good lost their contributions. In this experiment subjects do not incur losses if the threshold is not satisfied (effectively, a money-back guarantee).

2. Voluntary coalitions that provide a public good

This section presents a simple model of the formation of a voluntary coalition to provide a public good. The model is used to derive the equilibrium level of participation with a voluntary coalition which will be tested using the experiments described in section that follows. To begin, consider *n* identical players with payoff functions

$$\pi_i = A + b(q_i + q_{-i}) - cq_i, \qquad [1]$$

where q_i is equal to one if player *i* contributes to the public good and is zero if she does not, q_{-i} is the sum of the contributions by all other players, *b* is the constant marginal benefit of contributing to the public good, *c* is the cost of contributing, and *A* is a positive constant. The underlying structure of the players' interactions with each other is an *n*-player prisoners' dilemma, requiring $b < c$ and $nb > c$. That is, all players have a dominant strategy not to contribute to the public good in a noncooperative Nash equilibrium, but the players' joint payoffs will be maximized when they all contribute. Thus, players have an incentive to form a voluntary coalition as a means to increase contributions to the public good.

Following Ulph (2004) and Kolstad (2007), the formation of a voluntary coalition is modeled as a two-stage game. In the first stage, each player decides independently whether to join a coalition. Players make this decision while having complete information regarding the payoff functions of the other players and having perfect information regarding the choices made by the other *n* – 1 players. That is, each player is aware of how many players have joined the coalition, how many players have not joined the coalition and how many players have not yet decided. Once all players have made their decisions in the first stage, the second stage begins.

Let *s* denote the number of members to the coalition in the first stage. In the second stage, each member contributes to the public good provided that $s \geq s_{\min}$, where $s_{\min}$ is the minimum number of members required for individual contributions to the public good to be profitable (referred to as the *participation threshold*). If $s < s_{\min}$ after the first stage, then the members do not contribute to the public good in the second stage. The *n* - *s* players that did not join the coalition in the first stage maximize their individual payoffs by not contributing to the public good in the second stage, regardless of the number of members. The derivation of the participation threshold is discussed next.

Let $\pi^{m}(s)$ denote the payoff of each of the coalition members if they contribute to the public good, and let $\pi^{nm}(s)$ denote the payoff of each of the non-members. (The superscript *m* signals that the player in question is a member of the coalition, while the superscript *nm* signals that the player is not a member of the coalition). From [1], the payoff functions for each coalition member and for each non-member are:

$$\pi^{m}(s) = A + bs - c; \qquad [2]$$
$$\pi^{nm}(s) = A + bs.$$

Since $b < c$, $nb > c$, and $\pi^{m}(s)$ is increasing in *s*, there exist coalition sizes that are strictly greater than one and weakly less than *n* that are profitable. The smallest of these profitable coalitions is

$$s_{\min} = \min s \mid \pi^{m}(s) \geq \pi^{nm}(0) = \min s \mid s \geq c / b. \qquad [3]$$

Given $s_{\min}$ as the participation threshold, the equilibrium number of members that join the coalition in stage 1, s^{*}, is determined by adopting the well-known stability conditions often utilized in the context of cartels and international environmental agreements (e.g., D'Aspremont et al. 1983; Barrett 1994; Ulph 2004; Kolstad 2007). Following D'Aspremont et al. (1983), a coalition of members is considered stable if no member wants to leave the coalition (the coalition is *internally stable*) and no non-member wants to join the coalition (the coalition is *externally stable*).

It is easy to demonstrate that in this game the only internally and externally stable coalition is the smallest profitable coalition of coalition members, $s_{\min}$. To see why, note for coalitions where $s > s_{\min}$ at least one member could leave the coalition and the remaining members would still find it profitable to contribute to the public good.[5] Since individuals are motivated to leave a

---

[5] Using the payoff functions [2], the new non-member's payoff would then be $\pi^{nm}(s-1) = A + b(s-1)$, which is greater than its payoff if it stayed in the coalition, $\pi^{m}(s) = A + b(s) - c$, by the amount $c - b > 0$.

coalition of any size $s > s_{\min}$, these coalitions are not internally stable. On the other hand, a coalition of size $s = s_{\min}$ is internally stable, because if one member leaves the coalition it is no longer profitable for the remaining members to contribute to the public good. Since no individual would provide the public good in this case, a defector's payoff would simply be $\pi^{nm}(0) = A$, which is weakly less than its payoff if it stayed in the coalition, $\pi^{m}(s_{\min}) = A + bs_{\min} - c$. Finally, it is easy to show that all profitable coalitions are externally stable because $\pi^{m}(s+1) - \pi^{nm}(s) = b - c < 0$, which indicates that an individual who joins an already profitable coalition is worse off in comparison to staying out of the coalition altogether.

Since coalitions of size $s_{\min}$ are the only internally and externally stable coalitions, the subgame-perfect Nash equilibrium of the voluntary coalition game is that $s^{*} = s_{\min}$ players join the coalition in the first stage and make their contributions to the public good in the second. With discrete choices and identical players it is not possible to identify which of the s^{*} players will join the coalition and which players will not.[6] Therefore the subgame-perfect Nash equilibria in this game are defined by any set of s^{*} players joining the coalition in stage 1 (and thus contributing to the public good in stage 2) and n - s^{*} players not joining the coalition and not contributing to the public good.[7] However, because players make their decision whether to join a coalition in stage 1 consecutively while having real-time information about the choices made by the other n – 1 players, it is possible to further define the set of equilibria. Because free-riding non-members earn strictly higher payoffs than coalition members by the amount $\pi^{nm}(s) - \pi^{m}(s) = c$, opting out of the coalition is a Nash strategy for the first n - s^{*} decision makers. And because voluntary coalitions of size s^{*}, by definition, are profitable for the members even after the n – s^{*} players have decided to opt out of the coalition, the remaining s^{*} players maximize their earnings by joining the coalition.[8] Thus, in equilibrium the voluntary coalition

[6] Of course in the trivial case in which $s^{*} = n$, all players join the coalition in the first stage of the game.
[7] A formal derivation of the subgame-perfect Nash equilibria in a discrete choice threshold public goods game can be found in Isaac et al. (1989) and Asch et al. (1993).
[8] Others have derived similar equilibria using sequential threshold public good games. Erev and Rapoport (1990) and Cooper and Stockman (2002) analyze discrete-choice, sequential threshold

consists of s^* members each contributing to the public good and n-s^* free-riding non-members.

3. Experimental design

The following experiments were designed to empirically test the equilibria from the model in section two. Although results are reported on the frequency of trials in which the participation threshold was satisfied and the respective size of the voluntary coalitions, the focus of this research is on the relative timing of players' decisions in stage 1. That is, the primary interest is the extent to which subjects deliberately opt out of joining voluntary coalitions that are mutually beneficial but disproportionately benefit the free-riding non-members. Our two treatments vary in the cost of contributing to the public good, which in turn varies the participation threshold and the difference between member and non-member payoffs (i.e., the net benefits to free riding).

All of the experimental sessions were held in a computer lab at the [**insert university here**] using undergraduate and graduate students as subjects recruited from the general student population. In each of the treatments, subjects were in groups of n = 10 and made independent decisions at a computer station using software specifically designed for this project. Subjects were brought into the computer lab, seated and paid five dollars for arriving on time. During each session, two ten-subject groups were in the lab concurrently. The participants were each provided with a folder at their computer station containing their consent form and a set of instructions [**Included as a Reviewer's Appendix**]. The instructions were read aloud by the moderator and at times the subjects were asked to look up at a presentation in the front of the computer lab for further clarification of the instructions. After answering a series of practice questions, a 13 period experiment began.[9]

public good games in which players are assigned an order in which they must make the decision whether to contribute to a public good. A threshold is specified, referred to as the "minimum contributing set" (MCS), and if the MCS is satisfied the public good is provided and otherwise it is not. The subgame-perfect Nash equilibrium in these games is that the first $n - MCS$ players choose not to contribute to the public good and the remaining players choose to contribute.

[9] The subjects were not aware of the total number of periods until the experiment completed.

Each period, participants were randomly assigned to one of the two groups such that the same ten people were never in the same group more than once. The random assignment of groups was done to mitigate potential problems of reputation that can occur when the same people interact in a repeated setting (Andreoni and Croson 2002). Two sessions were conducted for each treatment, and therefore in total, four groups of ten subjects participated in 13 rounds each, resulting in 52 group-level observations and 520 individual-level observations per treatment. The earnings were reported in experimental dollars, and ten experimental dollars exchanged for one US dollar. Earnings in US dollars were paid in cash once the experiment was over.

To avoid potential biases subjects may have regarding the provision of public goods, and to generalize our experimental results to all applications of cooperative coalitions, a neutral frame was established for the context and language of the experiment. In these experiments participants made decisions regarding the production of an unspecified product. Specifically, they chose either to *agree to produce* one unit of an unspecified product, or *not agree to produce* the unit. An agreement was said to 'form' if enough subjects agreed to produce to satisfy the participation threshold. The groups were informed of the participation threshold at the beginning of the experiment.

While subjects were making their decisions whether to agree to produce they were provided with real-time information about the decisions made by the other nine members in the group. Specifically, they were informed about the number of other subjects that agreed to produce, the number of other subjects that did not agree to produce and the number of other subjects that had not yet made a decision. Also, in order to cap the length of each period, they only had 60 seconds to make their decisions. If someone failed to decide before the time was up, she chose *not agree to produce* by default.

The software was designed with a number of features to ensure that descisions were truly sequential. For example, if more than one subject made a decision within the same second, only the first decision was recorded. In those situations the subjects whose decision was blocked received a message informing them that their action was not recorded and instructed them that the group's information has changed. Those subjects could then reevaluate their position and were given the opportunity to make another decision. In addition, if a group member made a decision within the last five seconds of the round when there were still undecided subjects, an additional five seconds were added to the time remaining. This feature provided undecided subjects enough time to assimilate the changes before making their decision.

3.1 Low-cost of contributing to the public good

In this treatment, the individual cost of contributing to the public good, *c*, was low (relative to the treatment described in section 3.2). Throughout, this treatment is referred to as the *low-cost treatment*. Parameter values of $n = 10$, $A = 8$, $b = 3$ and $c = 7$ were chosen for equation [1], and by using [3] the participation threshold was $s_{\min} = 3$ (i.e., the next integer value greater than 7/3). Thus, if at least three of the ten group members agreed to produce, then the agreement formed and those individuals that agreed to produce automatically produced the product and earned $\pi^{m}(s)$ using [2]. If less than three group members agreed to produce, the entire group did not produce and each subject earned eight experimental dollars, i.e., $\pi^{nm}(0) = 8$. Those subjects that chose *not agree to produce*, did not produce the product regardless if the agreement formed and earned $\pi^{nm}(s)$ using [2].

Given our choice of parameters and the results of the model from section two, we can specify our first testable hypothesis.

Hypothesis 1: In the *low-cost treatment* the first $n – s^{*} = 7$ players to make a decision will choose to opt out of the voluntary coalition and the last $s^{*} = 3$ players to make a decision will choose to join the coalition.

Table 1 shows a subject’s payoff possibilities for each period of the experiment. This table was created by substituting the parameter values into the equations in [2], and its layout is consistent with the earnings table used by Rapoport and Eshed-Levy (1989). The boxes marked with an *X* indicate outcomes that are not possible because the participation threshold would not be met in those circumstances. To see why, note that the top row of Table 1 corresponds to the number of other subjects (ranging from 0 to 9) producing the good. However, production can only occur if at least three members of the group agree to produce. Therefore, under no circumstance could only one or two members of the group produce the good, and hence the *X*'s. All elements of the earnings table were clearly explained to the subjects before the experiment began.

[Insert Table 1 here]

3.2 High-cost of contributing to the public good

This treatment proceeded exactly as the *low-cost treatment* except that the cost of contributing to the public good, *c*, was increased from 7 to 15.08. Throughout, this treatment is referred to as the *high-cost treatment*. The parameter values were thus set at n = 10, A = 8, b = 3 and c = 15.08 and using equation [3], the participation threshold was derived at s = 6 (i.e., the next integer value greater than 15.08/3). If at least six of the ten group members agreed to produce, then the agreement formed and those individuals that agreed to produce automatically produced the product and earned $\pi^{m}(s)$ using [2]. If less than six group members agreed to produce, the entire group did not produce and each subject earned eight experimental dollars, i.e., $\pi^{nm}(0) = 8$. Those subjects that chose *not agree to produce*, did not produce the product regardless if the agreement formed and earned $\pi^{nm}(s)$ using [2]. Given our choice of parameters and the results of the model from section two, we have our second testable hypothesis.

Hypothesis 2: In the *high-cost treatment* the first $n - s^{*} = 4$ players to make a decision will choose to opt out of the voluntary coalition and the last $s^{*} = 6$ players to make a decision will choose to join the coalition.

Table 2 displays a subject's payoff possibilities for the *high-cost treatment*, where the *X*'s, as in Table 1, indicate impossible outcomes.[10] Table 3 summarizes the experimental design.

[Insert Table 2 here]

[Insert Table 3 here]

[10] The values in Table 2 are calculated earnings rounded to the nearest whole number.

4. Results

The summary statistics for this study are displayed in Table 4. As previous studies have found (Van de Kragt et al. 1983; Dawes et al. 1986; Suleimen and Rapoport 1992; Rapoport and Suleimen 1993; Cadsby and Maynes 1999), raising the participation threshold significantly reduces the frequency of public good provision. Specifically, coalitions formed in only 61.5% of trials when the participation threshold was six players, a significant decrease compared with 96.2% of trials when the participation threshold was three players (61.5% vs. 96.2%, $p < 0.001$)[11]. In aggregate, coalitions formed in 78.8% of all trials for this study. Previous studies on sequential threshold public goods report similar findings. Specifically, Erev and Rapoport (1990) report 66.7% of trials with successful provision, Goren, Kurzban and Rapoport (2003) report 88.9% and 81.8% successes from their two sessions and Goren, Rapoport and Kurzban (2004) report successes of 73.8% and 72.4% for their two sessions. When coalitions did form in our experiments there were more members in the *high-cost treatment* (6.19 vs. 4.26, $p < 0.001$), however, the average provision of the public good over all trials was not significantly different across the two treatments (3.81 vs. 4.10, $p = 0.243$).

[Insert Table 4]

The reason for such a significant decrease in the frequency of public good provision between the two treatments, however, is unclear. Previous studies have typically attributed this decrease in provision to the increased coordination problem among players caused by the higher threshold. However, with sequential decisions and perfect information regarding other players' decisions, the same coordination problems do not exist. This paper, to our knowledge, is the first to vary thresholds within sequential decision making environments with complete information. A second fundamental difference between this study and the studies referenced in the paragraph above is that the threshold in our experiments is not independent of the value of the public good. The higher threshold in this study is the result of the added cost of contributing to the public which in turn increases the minimum sized profitable coalition. Most importantly, a higher participation threshold corresponds to greater relative payoffs for the free riders. With the

[11] Unless otherwise noted, all paired-comparison statistical tests are Mann-Whitney rank-sum tests.

existing experimental design it is not possible to determine whether coalitions form less frequently in these experiments because more members are required to join in order to satisfy the participation threshold (e.g., a coordination problem) or because subjects are willing to suffer individual losses by deliberately causing a coalition not form in order to punish would-be free riders (e.g., *spiteful* behavior, as reported by Cason et al. (2004)).

The primary interest of this study is the relative order of players' participation decisions and the sequence as to which subjects, as predicted by the model in section two, opt out of voluntary coalitions that disproportionately benefit free riders. Recall from section two that, in theory, the first set of decisions made in each ten-person group will be made by players choosing not to join a coalition. As formulated in Hypothesis 1, in the *low-cost treatment* the first seven subjects are expected to opt out of the coalition and the last three subjects are expected to join the coalition. Likewise, from Hypothesis 2, in the *high-cost treatment* the first four players to make a decision are expected to opt out of the coalition forcing the remaining six players to maximize their individual earnings by joining the coalition.

In the *low-cost treatment* groups never played the Nash equilibrium exactly, and in the *high-cost treatment* exact equilibrium play occurred in only 4 trials. Therefore, our data reject the two testable hypotheses put forth in section three.[12] However, interesting comparisons can be made between the two treatments. First, from the last two columns in Table 4 it is clear that group behavior in the *high-cost treatment* was more consistent with the theoretical predictions. Column five in Table 4 lists the percentage of coalitions that formed with the predicted number of members; that is, three members in the *low-cost treatment* and six members in the *high-cost treatment*. When coalitions formed in the *low-cost treatment*, only 30% of those coalitions had the predicted level of participation while the remaining 70% had over participation. Contrast this with the results from the *high-cost treatment* in which 88% of effective coalitions had exactly the predicted six members (30% vs. 88%, $p < 0.001$).

---

[12] Tests of the equilibrium predictions from section two using Wilcoxon signed-rank tests result in $p < 0.001$ for both treatments. The sequential public goods study by Erev and Rapoport (1990), which is perhaps the study most related to ours, also finds a low occurrence of exact equilibrium play (3 of 15 trials). In their study players are in groups of five and must decide sequentially whether to contribute to the public good. If at least three players contribute, the public good is provided.

The last column in Table 4 lists the percentage of effective coalitions in which the last decision made was critical for a coalition to form. Recall, the model predicts that the last decision should always be critical. In the *low-cost treatment*, the participation threshold was satisfied after the last player made her decision in only 4% of the effective trials. In the *high-cost treatment*, on the other hand, the last person's decision was critical for a coalition to form in 41% of the trials (4% vs. 41%, $p < 0.001$). Together, the information in the last two columns in Table 4 suggests that although the precise order of subjects' decisions deviated from the theoretical prediction, group behavior in the *high-cost treatment* more closely resembled equilibrium play.

Figure 1 illustrates the most interesting result of this research. The graph shows that the higher participation threshold caused a complete reversal of the overall participation strategy implemented by the group members. The horizontal axis in Figure 1 lists the order of decisions made from one to ten, and the vertical axis displays the percentage of those decisions (52 in total for each rank order) in which subjects decided to opt out of the coalition. Although all ten decisions are listed on this graph, focus attention on the first four decision makers because in both the *high-cost* and *low-cost* treatments, the theoretical model from section two (and the hypotheses from section three) predicts that 100% of the first four decisions should be made by subjects opting out of the voluntary coalition.

[Insert Figure 1 here]

In the *low-cost treatment*, in contrast to Hypothesis 1, the majority of the first four decisions were made by subjects willing to join a coalition, not by subjects choosing to opt out. Moreover, the majority of the last six decisions in the *low-cost treatment* were made by subjects opting out of the coalition. Rather than subjects trying to be among the first to free ride as the model predicts, the results of this treatment are more consistent with the description of a hold-out strategy in which subjects with free-riding motives hold off on making a decision in hope that a subset of the other group members will fulfill the participation requirement (Swanson 1999, p.139; Barrett 2003, p. 142).

The opposite behavior is observed in the *high-cost treatment*. From Figure 1, at least 50% of each of the first four decision makers decided to opt out of the coalition.[13] This type of behavior is more consistent with the theoretical predictions from the model in section two relative to the *low-cost treatment*. The results show that when the participation threshold was increased, the participation strategy with voluntary coalitions changed completely and changed in line with the predictions from the theoretical model.

We estimated a logit regression model to provide a better understanding of how the increase in the cost of contributing to the public good, and thus an increase in the participation threshold, affected subjects' decisions to opt out of a voluntary coalition. Specifically, we estimate

$$\begin{aligned} Optout_{it} = \beta_0 + \beta_1 HC_i + \beta_2 HC*rank_i + \beta_3 LC*rank_i + \beta_4 HC*rank_i^2 \\ + \beta_5 LC*rank_i^2 + \beta_6 HC*\#mem_i + \beta_7 LC*\#mem_i + \psi_t + u_{it}, \end{aligned}$$

where *Optout* = 1 if subject *i* opted out of out of the coalition in period *t*, *HC* and *LC* indicate the high-cost and low-cost treatments respectively, *#mem* is the number of existing coalition members when subject *i* made her participation decision, *rank* is the endogenously determined order in which subject *i* made her decision (1 to 10) and ψ_t is a vector of *T*-1 dummies that capture potential period effects. A subject-specific random-effects specification of the error term ($u_{it} = \alpha_i + \varepsilon_{it}$, where α_i captures random effects and ε_{it} is the contemporaneous error term) was chosen in order to control for potentially strong individual effects that can occur when the same subject makes multiple decisions within a treatment. All observations in which opting out of the coalition was not a subject's Nash strategy were excluded (151 observations were excluded). Therefore, observations were dropped for two reasons; (1) when too many subjects had already opted out of the coalition so that the participation threshold could not be satisfied regardless of the subject's decision, and (2) when a subject's decision was critical in the sense that if she did not join the coalition the participation threshold could not have been satisfied.

In addition to including the treatment variable (*HC* = 1 for observations from the *high-cost treatment*), interaction variables were included to capture the potential influence the existing number of coalition members and the rank order of

[13] For each of the ordered decisions, the percentage of subjects opting out of the coalition is

decisions might have on a subject's decision to opt out of a coalition in each treatment. Two variables interacting the treatment with the number of existing coalition members were included (*HC*#mem* and *LC*#mem)*. In addition, the variables *HC*rank* and *LC*rank* interact the dummy variables for treatment with the order of the decision (ranging from 1st to 10th). Two variables interacting the treatment and rank squared (*HC*rank*^{2} and *LC*rank*^{2}*)* were also included to capture a potentially nonlinear relationship between the likelihood of opting out of a coalition and the order of the decision. In total, the regression includes data from 80 unique subjects with a total of 889 observations. The regression results are contained in Table 5.[14]

[Insert Table 5 here]

The significance of the dummy variable for the *high-cost treatment* ($p < 0.001$) indicates that increasing the cost of contributing to the public good, and thus increasing the participation threshold, significantly increases the likelihood of a subject opting out of a coalition. Recall, the dataset used to estimate the model is restricted to observations in which opting out is the theoretical prediction. Thus, subjects' behavior in the *high-cost treatment* more closely resembles the theoretical prediction of opting out relative to the *low-cost treatment*.

Matching the trend illustrated in Figure 1, the negative coefficient attached to *HC*rank* indicates that the early decision makers are more likely to opt out of a coalition relative to those later in the sequence ($p < 0.001$). However, as suggested by the coefficient for *HC*rank*^{2} , subjects' behavior in this treatment eventually switches with subjects becoming more likely to opt out of the coalition the later, relative to others, they make their decision ($p < 0.001$). This effect is illustrated in Figure 1 in which the line for the *high-cost treatment* bends upward around the sixth decision maker. In general these results suggest that although we would predict universal opting out at all ranks, we find the earliest and latest decision makers in the high-cost treatment are relatively more likely to opt out. On the other hand, in the *low-cost treatment* the rank order of decisions only has a significant influence on subjects' behavior for the later decision makers; that is,

---

significantly different ($p < 0.05$) between the *low-cost* and *high-cost* treatments.

[14] The subject-specific random effects were significant ($p < 0.001$) and the period-specific fixed effects were insiginificant ($p < 0.660$).

the later decision makers are more likely to opt out of voluntary coalition (p = 0.008).

Finally, the positive coefficients for *HC*#mem* and *LC*#mem* reveal that the likelihood of a subject joining a coalition increases in the number of coalition members (p = 0.020 and p =0.326, respectively), but this result is only significant for the *high-cost treatment*. Although the experiments were not designed to formally test for it, the results indicate a pattern of 'conditional cooperation' in which subjects act cooperatively (i.e., join a coalition) given that other members already have (Fischbacher et al. 2001; Frey and Meier 2004). Although this behavior is not predicted by the theoretical model, it is not entirely surprising due to previous experimental evidence. The results suggest that subjects are more likely to participate in voluntary coalitions if others have already committed to doing so.

5. Conclusion

These experiments investigate the extent to which subjects deliberately opt out of voluntary coalitions that provide public goods. Examining free-riding behavior in this setting is useful for guiding the design of effective participation rules that are part of all voluntary coalitions. The predictions from the theoretical model are salient; the first set of decisions will be made by subjects choosing to opt out of a voluntary coalition. Although our experimental analysis does not explore the reasons subjects behave as they do, in theory opting out early of a voluntary coalition forces those that remain into joining the coalition and contributing to the public good. By forcing others to join the coalition and contribute to the public good, free riders avoid paying the cost of contributing while enjoying the benefits.

Two scenarios were considered, one in which the participation threshold required three players to join a coalition (out of ten) and the other in which the participation threshold required six players. Although subjects did not behave as theoretically predicted, interesting comparisons can be made between our two treatments. The results show a complete reversal of the subjects' participation strategies between the two treatments. When the participation threshold was low, in contrast to what the theoretical model predicts, the majority of the first decisions were made by subjects willing to join a coalition. In contrast, when the participation threshold was high, the majority of the first decisions were made by subjects deliberately opting out of a voluntary coalition. The results from the treatment with the higher participation threshold were more consistent with the prediction that the first set of decisions will be made by subjects choosing to opt out of voluntary coalitions.

The reason for the complete reversal in behavior between the two scenarios, however, is unclear. One hypothesis stems from the fact that in our experiments the participation threshold increases in response to increases in the cost of contributing to the public good, and therefore a higher participation threshold results in higher relative payoffs going to free riders. As a result, subjects may have greater incentives to opt out of coalitions with higher participation thresholds. The increasing gap between member and non-member payoffs may also explain the low frequency of trials in which the participation threshold was satisfied when the threshold was raised. It is possible that subjects were willing to sacrifice individual earnings by preventing coalitions from forming in order to prevent potential non-members from free riding (e.g., *spiteful* behavior).

Our experiments provide insights on the sequence of participation and the overall provision of a public good when agents have the opportunity to deliberately opt out of a voluntary coalition. Further research is required to rigorously address the question of why subjects reversed their participation strategies when the participation threshold was raised. As a starting point, it would be interesting to observe if subjects change their behavior when the participation threshold is raised while keeping the difference between member and non-member payoffs constant. This treatment would allow the researcher to separate the effect of an increase in the relative payoffs going to free riders from an increase in the participation threshold. In addition, variation in the participation threshold should be explored as it is possible that behavior may change when the threshold is set closer to the extremes. There are a number of scenarios that need to be explored in order to determine what influences the decision to participate, or deliberately not participate, in a voluntary coalition.

References

Andreoni, James and Rachel Croson, 2002. “Partners Versus Strangers: Random Rematching in Public Goods Experiments.” Prepared for *Handbook of Experimental Economics Results*, edited by Charles R. Plott and Vernon L. Smith.

Asch, Peter; Giglitotti, Gary A. and James A. Polito, 1993. “Free Riding with Discrete and Continuous Public Goods: Some Experimental Evidence.” *Public Choice* 77(): 293-305.

Bagnoli, Mark and Michael McKee, 1991. “Voluntary Contributions Games: Efficient Private Provision of Public Goods.” *Economic Inquiry* April: 351-366.

Barrett, Scott. 1994. "Self-Enforcing International Environmental Agreements." *Oxford Economic Papers* 46(5): 878-94.

Barrett, Scott, 2003. *Environment and Statecraft: The Strategy of Environmental Treaty-Making*. Oxford and New York: Oxford University Press.

Cadsby, Charles Bram; Croson, Rachel; Melanie Marks and Elizabeth Maynes. 2008. “Step Return Versus Net Reward in the Voluntary Provision of a Threshold Public Good: An Adversarial Collaboration.” *Public Choice* 135(3): 277-289.

Cadsby, Charles Bram and Elizabeth Maynes, 1999. “Voluntary Provision of Threshold Public Goods with Continuous Contributions: Experimental Evidence.” *Journal of Public Economics* 71(1): 53-73.

Cason, Timothy N.; Saijo, Tatsuyoshi; Yamato, Takehiko and Konomu Yokotani, 2004. “Non-Excludable Public Goods Experiments.” *Games and Economic Behavior* 49(1): 81-102.

Coats, Jennifer C. and William S. Neilson. 2005. “Beliefs about Other-Regarding Preferences in a Sequential Public Goods Game.” *Economic Inquiry* 43(3): 614-622.

Cooper, David J. and Carol Kraker Stockman, 2002. “Learning to Punish: Experimental Evidence from a Sequential Step-Level Public Goods Game.” *Experimental Economics* 5(1): 39-51.

Croci, Edoardo. 2005. “The Economics of Environmental Voluntary Agreements.” *The Handbook of Environmental Voluntary Agreements*. Edoardo Croci (ed.), Dordrecht and New York: Springer.

D’Aspremont, Claude; Alexis Jacquemin; Jean Jaskold Gabszewicz and John A. Weymark. 1983. “On the Stability of Collusive Price Leadership.” *The Canadian Journal of Economics* 16(1): 17-25.

Dawes, Robyn M.; Orbell, John M.; Simmons, Randy T. and Alphons J. C. Van De Kragt. 1986. “Organizing Groups for Collective Action.” *The American Political Science Review* 80(4): 1171-1185.

Dorsey, Robert E. 1992. “The Voluntary Contributions Mechanism with Real Time Revisions.” *Public Choice* 73(3):261-282.

Erev, Ido and Amnon Rapoport, 1990. “Provision of Step-Level Public Goods: The Sequential Contribution Mechanism.” *Journal of Conflict Resolution* 34(3): 401-425.

Fischbacher, Urs; Gachter, Simon and Ernst Fehr, 2001. “Are people conditionally cooperative? Evidence from a public goods experiment.” *Economics Letters* 71(3): 397 – 404.

Frey, Bruno S. and Stephan Meier, 2004. “Social Comparisons and Pro-social Behavior: Testing “Conditional Cooperation” in a Field Experiment.” *The American Economic Review* 94(5):1717 – 1722.

Goren, Harel; Robert Kurzban and Amnon Rapoport. 2003. “Social Loafing vs. Social Enhancement: Public Goods Provisioning in Real-Time with Irrevocable Commitments.” *Organizational Behavior and Human Decision Processes* 90(2):277-290.

Goren, Harel; Amnon Rapoport and Robert Kurzban. 2004. “Revocable Commitments to Public Goods Provision under the Real-Time Protocol of Play.” *Journal of Behavioral Decision Making* 17(1):17-37.

Isaac, Mark R.; David Schmidtz and James M. Walker, 1989. “The Assurance Problem in a Laboratory Market.” *Public Choice* 62(3): 217-236.

Kolstad, Charles, 2007. “Systematic Uncertainty in Self-Enforcing International Environmental Agreements.” *Journal of Environmental Economics and Management* 53(1): 68-79.

Kurzban, Robert; McCabe, Kevin; Vernon L. Smith and Bart J. Wilson. 2001. “Incremental Commitment and Reciprocity in a Real-Time Public Goods Game.” *Personality and Social Psychology Bulletin* 27(12): 1662-1673.

Marks, Melanie and Rachel Croson, 1998. "Alternative Rebate Rules in the Provision of a Threshold Public Good: An Experimental Investigation." *Journal of Public Economics* 67(2): 195-220.

Morgenstern, Richard D. and William A. Pizer, 2007. *Reality Check: The Nature and Performance of Voluntary Environmental Programs in the United States, Europe, and Japan*. Washington D.C.: Resources for the Future Press.

Rapoport, Amnon and D. Eshed-Levy, 1989. "Provision of Step-Level Public Goods: Effects of Greed and Fear of Being Gypped." *Organizational Behavior and Human Decision Processes* 44(3): 325-344.

Rapoport, Amnon and Ramzi Suleiman, 1993. "Incremental Contribution in Step-Level Public Goods Games with Asymmetric Players." *Organizational Behavior and Human Decision Processes* 55(2): 171-194.

Rondeau, Daniel; Schulze, William D. and Gregory L Poe, 1999. "Voluntary Revelation of the Demand for Public Goods Using a Provision Point Mechanism." *Journal of Public Economics*. 72(3), 455-470.

Suleiman, Ramzi and Amnon Rapoport, 1992. "Provision of Step-Level Public Goods with Continuous Contribution." *Journal of Behavioral Decision Making* 5(2): 133-153.

Swanson, Timothy and Sam Johnston, 1999. *Global Environmental Problems and International Environmental Agreements: The Economics of International Institution Building*. Cheltenham, U.K and Northampton, MA: Edward Elgar.

Ulph, Alistair, 2004. "Stable International Environmental Agreements with a Stock Pollutant, Uncertainty and Learning." *Journal of Risk and Uncertainty* 29(1): 53-73.

United Nations Framework Convention on Climate Change, 1998. "*Kyoto Protocol to the United Nations Framework Convention on Climate Change*."

Van de Kragt, Alphons J. C.; Orbell, John M. and Robyn M. Dawes, 1983. "The Minimal Contributing Set as a Solution to Public Goods Problems." *The American Political Science Review* 77(1): 112-122.

Tables

Table 1: Earnings table for the *low-cost treatment*

# of OTHER players that PRODUCE	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
YOUR earnings if you produce	X	X	$10	$13	$16	$19	$22	$25	$28	$31
YOUR earnings if you don't produce	$8	X	X	$17	$20	$23	$26	$29	$32	$35

Table 2: Earnings table for the *high-cost treatment*

# of OTHER players that PRODUCE	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
YOUR earnings if you produce	X	X	X	X	X	$11	$14	$17	$20	$23
YOUR earnings if you don't produce	$8	X	X	X	X	X	$26	$29	$32	$35

Table 3: Experiment design summary

Treatment	**Participation Threshold**	**Number of Subjects (group size *n* = 10)**	**Number of Group Observations (13 periods)**	**Number of Individual Observations**
Low-cost treatment	3	40	52	520
High-cost treatment	6	40	52	520
Totals		**80**	**104**	**1,040**

Table 4: Coalition formation and public good provision

Treatment	Percent of trials in which a coalition formed	Average provision of the public good	Average number of members when coalitions formed	Percent of coalitions in which the threshold was *exactly* met	Percent of coalitions that formed on the last (10^{th}) decision
Low-cost treatment (Threshold = 3)	96.2 (2.69) [52]	4.10 (0.197) [52]	4.26 (.160) [50]	30% (6.55) [50]	4% (2.80) [50]
High-cost treatment (Threshold = 6)	61.5 (6.81) [52]	3.81 (0.425) [52]	6.19 (.095) [32]	88% (5.94) [32]	41% (8.82) [32]

Standard errors in parentheses, number of observations in brackets.

Table 5: Effect of treatment, rank order of decision and number of coalition members on the decision to opt out of a voluntary coalition

	coefficient	*standard error*	*p-value*
constant	-1.7943	0.6630	0.007
HC	4.4699	0.8716	0.000
*LC*rank*	-0.0266	0.2649	0.920
*HC*rank*	-1.4045	0.2646	0.000
$LC*rank^2$	0.0616	0.0234	0.008
$HC*rank^2$	0.1451	0.0250	0.000
LC# mem*	-0.1154	0.1174	0.326
HC# mem*	-0.3449	0.1478	0.020
χ^2	129.03 ($p < 0.001$)		
n	889		

Notes: The dependent variable is the binary decision whether to opt out of the coalition (1 = *not agree to produce*, 0 = *agree to produce*).

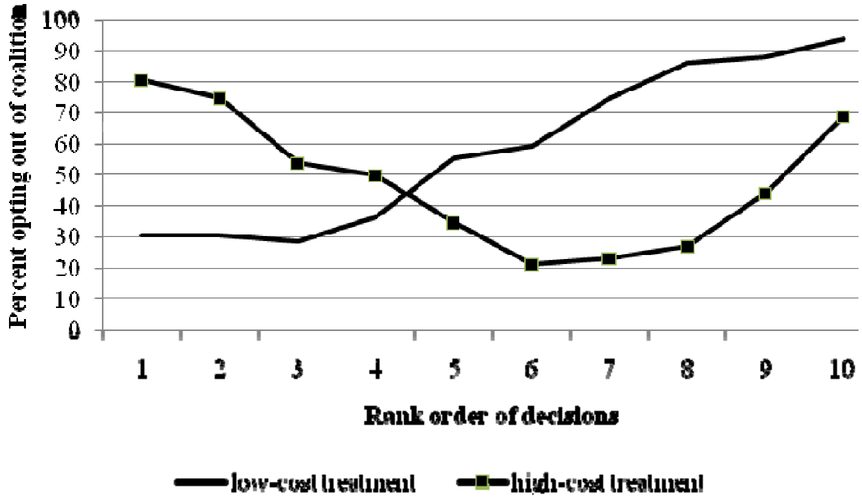

Figure 1: Percent of ordered decisions in which subjects chose to opt out of a voluntary coalition