MINIMUM PARTICIPATION CLAUSES AS EXCLUSION MECHANISMS IN PUBLIC GOOD AGREEMENTS

ANNA A. KLIS
DEPARTMENT OF ECONOMICS, UNIVERSITY OF TEXAS AT AUSTIN

Abstract. Many public goods treaties have minimum participation (MP) constraints. This paper analyzes the effects of heterogeneity on the chosen MP clause of public goods treaties. In the presence of heterogeneity, MP constraints can serve as a selection device in creating a more homogeneous group. There exist situations in which coalitions are restricted in what they can bargain over, and then the exclusion of certain agents from a treaty is Pareto improving. This paper gives a general set of sufficient conditions for the exclusion result to hold and presents an example of when exclusion improves upon unanimity.

Key words: game theory, public good provision, heterogeneity, agreements, coalitions, negotiations, minimum participation constraint, second best

PRELIMINARY DRAFT

1. Introduction

In an environment with sovereign agents, lack of external enforcement means that only mutually beneficial agreements can help reduce a negative externality or produce a positive externality: countries negotiate multilateral environmental agreements (MEAs) to reduce trans-boundary pollution; neighbors create Home Owner's Associations to reduce annoying actions amongst the community; the European Union originally had the aim of ending "frequent and bloody wars between neighbors"; and the euro currency zone was partly intended to limit the impact of currency exchange on trade deals and tourist experiences [1, 2]. Though these examples occur on different scales, in each situation the agents agree to bind themselves to the group's chosen action. The formation of government operates along similar lines: the government provides a positive externality, and the founding agents must determine how to establish and provide for it [13].

Agents must design their own agreement and endogenously determine the actions taken and the self-enforcement mechanisms used. Of particular interest are minimum participation (MP) clauses, mechanisms with low transaction costs and powerful benefits. MP clauses raise the value of an agreement by guaranteeing at least a certain number of compliers or a certain level of provision if the treaty is implemented. In a historical example, an MP clause applied to the Constitution of the United States: nine of the original thirteen colonies had to ratify the Constitution before it would take effect. Though widespread in modern agreements, MPs are especially prevalent in MEAs, such as the Montreal Protocol, the Rotterdam Convention, and the Kyoto Protocol. There are still many situations with externalities which could benefit from treaties, particularly negative externalities like overfishing, nuclear armament, and financial brinkmanship.

This paper examines how heterogeneity of agents affects the MP constraint chosen for treaties of varying ability to commit to certain types of group action. In particular, it evaluates how agreement actions which are restricted to egalitarianism, or equal action among players, lead to the desirability of using the MP constraint as an exclusion device. I find that under restricted actions, when a MP constraint can reduce the heterogeneity of the potential signatories, the mechanism can deliver more effective treaties.

Date: August 10, 2012 (first draft), updated July 4, 2014.
2225 Speedway Stop C3100, Austin, TX 78712-1690
aklis@utexas.edu
I thank my adviser, Dr. Maxwell Stinchcombe, for his teaching and his immense help in the development of this paper. I also thank Dale Stahl and participants in UT's theory writing seminars for their thoughts on improvement, and I thank Ina Taneva, Melinda Petre, Matthew Shapiro, as well as my parents, for their valuable assistance in the early editing process.

In Section 2, I present a brief development of this topic in others' work, as well as a summary of how this paper fits into the existing literature. In Section 3, I present a motivating example, while in Section 4, I set up my general model, present results, and then evaluate a parametrized model in more detail. In Section 5, I present conclusions and possible extensions. All proofs are in the Appendix.

2. Literature

Wagner's survey of self-enforcement in MEAs [17] examines multiple aspects of public good provision: a basic emissions reduction game, models of treaty negotiation, and dynamic extensions of both. The paper enumerates provisions that enhance participation in agreements such as minimum participation clauses, transfers, and issue linkages. The survey emphasizes a paper by Barrett [5] which presents a two models of self-enforcing agreements, one which is a one-shot game, while the other is infinitely repeated. Through numerical analysis, Barrett shows that, in the one-shot game, the self-enforcement strategy of punishment and reward may not sustain a larger group, even when the benefit of the agreement would be high. In the repeated game, credible tit-for-tat and trigger strategies can increase the number of participants, but the remaining characteristics of the model require parametrization, and the treaty may not be renegotiation-proof. While a repeated game may be more effective in capturing the long-standing interactions of nations, businesses, or home owners, the one-shot game can accurately represent the incentives present in the process of writing and signing a treaty, while preserving the endogeneity of decisions.

Black, Levi, and de Meza [7] examine the inclusion of an exogenous MP constraint into a one-shot game. They find that any constraint larger than the resulting number of members under open participation outperforms a standard agreement by increasing the number of participants and lowering the total stock of the negative action. Imposing the optimal participation level leads to maximal benefit of the treaty. Black et al. also discuss the issue of model timing, observing that in a single round of negotiation, agents will choose to implement the optimal participation level, but multiple rounds of negotiation lead to a decreased incentive to quickly ratify.

Carraro, Marchiori, and Oreffice [8] endogenize the MP constraint in a three-stage game of public good provision in which all agents are identical. The first stage is the minimum participation stage, in which all agents unanimously vote on the fraction required to sign the treaty in order for it to go into force. The second stage is the coalition stage, when each agent weighs the utility of being a member versus that of being a free-rider in deciding whether or not to join. The final stage is the policy stage, in which the coalition chooses its allocations cooperatively while non-members choose their actions non-cooperatively. Carraro et al. show that it is possible for agents to agree to an endogenously chosen MP clause which increases the overall number of signatories. However, the analysis is sensitive to the assumption of homogeneity of agents.

While arguments using homogeneity may apply to certain situations, in other applications, heterogeneity of agents is crucial to the analysis. Agents can vary among the benefit they receive from their individual action and effects caused by others' actions, while treaties can vary in type of committed action. For environmental considerations, any number of factors such as population, area, topology, GDP, and political relations have been shown to affect a country's decision to sign an MEA [6, 11, 14]. A source of heterogeneity among countries for the issues of nuclear disarmament or overfishing is that of technology; a country on the cutting edge of technology likely has lower costs of reduction of missiles or fishing compared to a country with little research and development. On a smaller level, when establishing an Home Owner's Association, families may value different restrictions, such as noise control or cleanliness of public areas, than do single households.

Wiekard, Wangler, and Freytag [18] add heterogeneity to Carraro et al.'s model. With heterogeneity, they change the minimum participation constraint from number of signatories to minimum abatement, which has some precedence in actual treaties. They find that free-riding almost always occurs, and that the inclusion of a random selection of a proposing agents emphasizes the importance of agenda setting. They continue to use the same coalition formation timing as Carraro et al.

Using a coalition formation model to represent a treaty negotiation process separates the MP and allocation decision. The MP is essentially decided on ahead of time by stating a minimal coalition, while the allocation decision is set by the formed coalition. This may be the proper timing for certain applications such as the home owner's association or even the euro zone. Since in most multilateral international negotiations regarding any topic, countries write agreements over a period of time and then vote on all final provisions in one shot, this may limit the theoretical predictions and outside relevance of the model. However, the coalition formation model can be regarded as robust to both timing scenarios because of backwards induction. Forward-looking agents will only suggest or sign agreements which benefit them in some way and which will gain acceptance from other agents.

In this context of coalition formation, I study a one-shot negative externality game and the set of agreements that improve upon Nash equilibrium. This can be understood as an equilibrium in a repeated game with Nash reversion, though I don't develop that idea or pursue the enforcement of agreements. In order to alleviate the negative externality, it is as if agents were embedded in a larger game of international politics and respect for negotiations, and they play a one-shot treaty coalition game.

In particular, I consider a specific type of treaty action. I develop a solution concept under limited commitment power, where coalition members can only commit to one-dimensional decreases from the Nash equilibrium. Under this *constraint* of egalitarianism I examine the MP choices that give the most improvement over Nash equilibrium.

When the possible agreement sets for a coalition are unrestricted, participants can always benefit from the reduction of the amount of free-riders, since each free-rider can contribute a bit more of the public good and improve the utility of all participants. Though it adds a restriction, egalitarianism can be a desirable treaty trait. A restriction in which all coalition members must take the same action allows for simplicity (the choice variable can be one-dimensional, instead of multi-dimensional) and may result from environments with uncertainty.

In a dynamic externality reduction game with private cost shocks, Harrison and Lagunoff [15] find that truth-telling and coalition participation in a first period negotiation require fully compressed quotas throughout the rest of the game: regardless of later shocks, all agents have the same per-period quota as other agents. Bagwell [4] develops a bilateral tariff negotiation game with both a one-shot static and dynamic version. He finds pooling equilibria in which countries with one type of public opinion will imitate the other type, both negotiating the same levels of tariffs. Though my model does not have uncertainty, these results add to the motivation of understanding the use of egalitarian treaties.

Under heterogeneity, restricting the action set of the commitment group changes the types of treaties enacted. The main result of this paper is that under one form of egalitarianism and given sufficient heterogeneity, the optimal MP constraint is strictly smaller than whole. The constraint – which may declare the number of players, the exact set of players, or the total action required – removes agents who are limited by lack of ability to greatly affect the public good by rendering them non-pivotal. The remaining agents, whose actions have the largest effects on public goods, create a more effective agreement.

Ludema and Mayda [16] find a similar exclusion result in their paper on tariff negotiations within the World Trade Organization using equal-split kinds of arguments. They find that when most-favored nation status results in an externality, only large exporters will participate in negotiations and offer equal-sized transfers to an importer in return for a lower tariff on a good. Small exporters of the good are unable to pay the transfer and thus free-ride on the eventual negotiation. I confirm this result in a broader setting of coalitions with restricted actions.

3. Motivational Example

In this section, I provide a numerical example with three agents. I hold one agent's utility function fixed, and then I calculate the Nash equilibrium actions as the values of the parameters of the other two agents vary. I then calculate the utility of lump-sum reduction treaties and determine which improve upon the Nash equilibrium. I present this example as a motivation for the exclusion of the "odd man out," showing that a treaty may be more successful when participants are more homogeneous.

A treaty's purpose is to bring the Nash equilibrium closer to the Pareto optimal solution through players' joint reductions. As in any public good agreement, there is concern that agents will prefer to free-ride, lowering the value of cooperation and collapsing the agreement. A minimum participation constraint can add initial value to a treaty and gain commitment from players.

An MP constraint could specify the exact minimum set players who must join the treaty in order for it to go into effect. The constraint could also be the required cardinality of the final set of participants, which is more akin to real treaty MP constraints, or the total sum of participants' actions, so that the treaty is not in force until the required level of action is committed. By agreeing on the set of agents J (or the number of agents $\#J$ or amount of reduction) as a measure of minimum participation, players can then infer the vector of commitments, which follow from the type of action restriction and J itself.

This constraint can also serve as a way of selecting a homogeneous group out of a set of heterogeneous agents. To examine this idea, I consider specific parametrized example in a simple world of three agents, $I = \{1, 2, 3\}$, to better understand the selection of J. The chosen utility function is:

$$u_i(a) = (1 - \theta_i)a_i - a_i \sum_{j=1}^{3} w_j a_j, \tag{1}$$

where a_i comes from $A_i = [0, 1]$ and the weights sum to one, i.e. $\sum_{j \in I} w_j = 1$. This utility function is twice-continuously differentiable, increasing and concave in a_i, and exhibits a negative externality from a_j, $j \neq i$. It has an unique equilibrium on A for each parameter set $\{\theta_1, \theta_2, \theta_3\}$.

I solve first for the optimal Nash actions and then determine how actions would change under each possible minimum participation constraint. Afterward, I consider which MP constraint defines the optimal coalition for a range of parameters.

3.1. **Nash Equilibrium.** In the absence of a treaty, each agent solves the following problem:

$$\begin{aligned} &\max_{a_i} \ (1 - \theta_i)a_i - a_i \sum_{k=1}^{3} w_k a_k, \\ &\text{s.t. } a_i \in [0, 1]. \end{aligned} \tag{2}$$

Thus, the Nash equilibrium actions are:

$$a_i^*\left(a_j^*(\theta_j), a_k^*(\theta_k)\right) \equiv \begin{cases} 1 & \text{if } (1-\theta_i) \geq 2w_i + w_j a_j^* + w_k a_k^* \\ \frac{1-\theta_i - w_j a_j^* - w_k a_k^*}{2w_i} & \text{if } w_j a_j^* + w_k a_k^* < (1-\theta_i) < 2w_i + w_j a_j^* + w_k a_k^* \\ 0 & \text{if } (1-\theta_i) \leq w_j a_j^* + w_k a_k^*. \end{cases}$$

For agents with relatively small values of θ, the solution is greater than zero; I call these agents "positive producers." Meanwhile, agents with rather large values of θ take action of zero; I call these agents "non-producers."

Unless all agents are non-producers, the Nash equilibrium is not optimal because of the negative externality. The equal-weighted Pareto optimal solution for the economy would result in prescribing reduced actions for each agent, in the form of:

$$a_i^P\left(a_j^*(\theta_j), a_k^*(\theta_k)\right) \equiv \begin{cases} 1 & \text{if } (1-\theta_i) \geq 2w_i + (w_i + w_j)a_j^* + (w_i + w_k)a_k^* \equiv \underline{\gamma}_i \\ \frac{1-\theta_i - (w_i+w_j)a_j^* - (w_i+w_k)a_k^*}{2w_i} & \text{if } \in (\underline{\gamma}_i, \overline{\gamma}_i) \\ 0 & \text{if } (1-\theta_i) \leq (w_i + w_j)a_j^* + (w_i + w_k)a_k^* \equiv \overline{\gamma}_i. \end{cases}$$

As expected, when the negative externalities are considered, each positive producer's action is reduced. The threshold of producing more is now higher, meaning the parameter must be very small, giving a large individual benefit, in order for an agent to be allowed to inflict a high level of externality on the others.

3.2. **Treaty Actions.** Here I consider a lump-sum reduction treaty, wherein each participant of the treaty reduces from his Nash equilibrium action by the same amount. Notationally, the lump-sum restricted treaty consists of $(J, a^{LS}(J))$, where the minimum participation constraint is J, or the minimum set of agents that must participate in the treaty, and $a^{LS}(J)$ are the agreed-upon actions.

Consider a treaty negotiation scenario as follows: all agents first determine J, then agents joining J choose $a^{LS}(u, J)$ while agents outside best respond. The Nash equilibrium is either the sovereign equilibrium or a vector of commitments and singleton responses. I examine incentives for agents to cooperate with a lump-sum reduction treaty under each possible cardinality MP constraint, $\#J \in \{0, 1, 2, 3\}$.

First, it is important to establish that, regardless of J, a non-producer cannot commit to a lump-sum reduction because it is impossible for him to reduce beyond zero. This distinct pattern of heterogeneity demonstrates how a smaller group can improve upon the results of the whole coalition, simply because with even one non-producer, the coalition of the whole can do nothing. This pattern extends beyond the trivial case of excluding agents with corner solutions of zero and holds even for strictly interior Nash equilibria.

In brief contrast, consider a different type of treaty, one of proportional reduction, where each agent participating has to reduce form Nash equilibrium by the same percentage. It is possible for a non-producer to commit to an egalitarian proportional reduction treaty, as any factor multiplying zero is still zero. Such an action is largely symbolic. Two or three agents could easily enter an ineffective proportional treaty, either by choosing a reduction of zero percent or, if all are non-producers, choosing any reduction level at all. Since such a treaty does not actually require positive reduction, there is no improvement over the Nash equilibrium. Effective treaties of this type are possible, but are not examined here; the focus will remain on lump-sum reduction treaties.

In addition, it is necessary to discuss the possible distribution of agents from the two types, positive producers and non-producers. Clearly, in a world of solely non-producers, there is no negative externality, no need for improvement, and hence, no need for a treaty. Thus, in the following discussion, I assume there is at least one agent who is a positive producer. Since there are only four meaningful cardinality MP constraints in a world of three, each can be examined in detail for optimal actions and implications.

Under the open membership rule of $J = \emptyset$ or $\#J = 0$, reductions can be negotiated, but there is no minimum number of members for the treaty to go into effect. Without repeated interaction providing a chance for punishment or without some sort of side transfers, open membership removes the initial value that the MP constraint could provide. Thus, no positive producer will join such a treaty in this game, unable to count on the participation of others, and the solutions are the same as under no treaty.

Under a singleton MP constraint or $\#J = 1$, if one agent considers committing to reduce on his own, he does not have to negotiate the amount − he would simply choose it. A non-producer could individually commit to an ineffective *proportional* reduction treaty of any level; even though this is an equilibrium in which an "agreement" arises, the total externality is not reduced from the Nash equilibrium level in any sense. Such an agent could not commit to a unilateral *lump-sum* reduction. A positive producer could reduce his action for the benefit of the whole, but such an action would run counter to the Nash equilibrium. Reducing unilaterally gives no outside benefit to the agent in question and allows all the other players to free-ride on the reduced action. Therefore an Pareto-improving treaty will not occur for the singleton MP clause.

For an MP constraint greater than one, there are Pareto-improving lump-sum reduction treaties possible. The question which sparks the most interest is when a treaty with an MP clause of two producers is preferred to one with a clause specifying all three must participate.

Without loss of generality, agents $i = 1, 2$ consider the treaty:

$$(J = \{1, 2\}, \{a_1^* - r^*(\{1, 2\}), a_2^* - r^*(\{1, 2\})\}),$$

while agent 3 best responds with a_3^*. The reduction $r^*(\{1, 2\})$ solves the following:

$$\max_{r \in \mathbb{R}} \sum_{i=1}^{2} \left\{ (1 - \theta_i)(a_i^* - r) - (a_i^* - r) \left[\sum_{k=1}^{2} \Big(w_k(a_k^* - r) \Big) + w_3 a_3^* \right] \right\}, \quad \text{s.t. } 0 \leq r \leq \min\{a_1^*, a_2^*\}. \tag{3}$$

With the MP commitment device, the coalition does not go into effect if one of the agents does not sign. Thus, 1 and 2 know each of them must sign in order for the other to uphold the agreement. Such an agreement would only be signed if the agents' individual utilities are improved. To see if signing is beneficial, the comparative statics of the individual utility from this treaty can be examined. Define u_i^{LS} for $i = 1, 2$ as the utility of entering into the treaty described, where 1 and 2 reduce and 3 best responds. Moving from zero, the marginal utility of increasing the reduction is:

$$\left. \frac{\partial u_i^{LS}}{\partial r} \right|_{r=0} = -(1 - \theta_i) + (2w_1 + w_2)a_1^* + w_2 a_2^* + w_3 a_3^*. \tag{4}$$

This statement is positive when the benefits of reduction outweigh the foregone benefits of action, i.e. when $(2w_1 + w_2)a_1^* + w_2 a_2^* + w_3 a_3^* > (1 - \theta_i)$. If Nash actions are large or θ_i is large, then this statement likely holds.

Expanding the treaty to include full participation requires that all three agents agree on the action vector. A Pareto-improving lump-sum reduction treaty for the coalition of the whole could only occur if all agents are positive producers. Similarly as with two, three positive producers do have an incentive to join a lump-sum reduction treaty.

3.3. **Treaty Negotiation and Equilibrium Selection.** To summarize, in a world of three agents, there are two possibilities for lump-sum reductions:

(1) Agents could sign an ineffective treaty, one where the chose reduction is zero. Any MP constraint is possible for this.
(2) Agents can sign a Pareto-improving lump-sum reduction treaties with with $r > 0$. The MP constraint must be greater than two for this case.

The category of Pareto-improving agreements deserves further examination, particularly which MP constraint and vector of actions will result.

Selection is an issue, which may be resolved through game timing or bargaining protocol. First, consider the issue of timing. A timing common to most models of the literature, such as that of Carraro et al., is one where the treaty participants determine their action as a coalition in a separate stage from all agents' decision of the MP constraint [8]. This timing reflects the idea that it would unfair or perhaps even infeasible to bind participants to the decisions of the whole group. In this timing, equilibrium selection proceeds according to the coalition's maximization function – the coalition that results from a first stage will choose its actions. The reduction can be chosen by maximizing the summed utility of the coalition members or some other function of member utility.

On the other hand, in many real-world agreements, all persons in attendance at the start of the negotiation have a say in the provisions of the agreement; only once these are agreed upon do agents declare their participation. However, the results of this timing is not so different: there are more possible equilibria without the coalition utility function to act as an equilibrium selector, but the equilibria are bound by the preferences of the expected participants. An agent who will not participate cannot suggest an unreasonable reduction and expect the other agents to join the treaty.

For instance, without a negotiation process more detailed than a unanimous vote, the possible equilibria lie on a continuum. Any of the valid values could be chosen in equilibrium through unanimous vote and adhered to in the policy stage of either timing, so the question of equilibrium selection for non-coalition timing persists. Most real-world agreements undergo rounds of discussion, as in many bargaining protocols. Any bargaining process which is strictly increasing and always efficient, such as Nash or Kalai-Smorodinsky, would result in an efficient selection. Furthermore, in some processes, as the bargaining set gets larger, everyone is better off. Thus, agents with a larger initial action give more room for the bargaining process, are able to reduce more, and can improve social welfare more.

Apart from heterogeneity in utility, agents may also have heterogeneity of bargaining power. A measure of bargaining power in multi-state agreements could be calibrated to various instruments of power, such as overall pollution rank, number of trade agreements, GDP, or United Nations Security Council membership, to name a few. Coalition negotiation captures the weakened position of an agent who has little to bring to the table, while other decision protocols such as unanimous voting may allow a small player to derail a treaty. This is an area for further research.

In the numerical example using the utility function Equation 1, if selection proceeds according to highest total utility, then the figure below gives a graph of which outcome will occur under which realization of parameters. In this figure, agent 1's parameter θ_1 is normalized to zero. The externality weights w_i are equal to one third for each agent. For each pair θ_2 and θ_3, I examine which treaties improve the most upon the Nash equilibrium, if any, out of the possible coalitions.

For small values of the parameter, there is an area the lower left-hand corner near the origin where the coalition of the whole is restriction Pareto optimal. However, as soon as the parameter value increases too much, that agent drops production to zero, rendering the coalition of the whole no longer optimal. There is an intermediate value for the parameter where the Nash equilibrium persists, since an a treaty among the two remaining producers would be sabotaged by an increase in action from the excluded player. However, once the parameter is large enough, the remaining two form the exclusive treaty. Along the x-axis, as player 2's value of θ_2 increases, there is a region where the optimal lump-sum restricted coalition is between players 1 and 3, represented in the lower right corner. Symmetrically, along the y-axis, there is a region where the optimal coalition is between players 1 and 2, represented in the upper left corner. The remaining region,

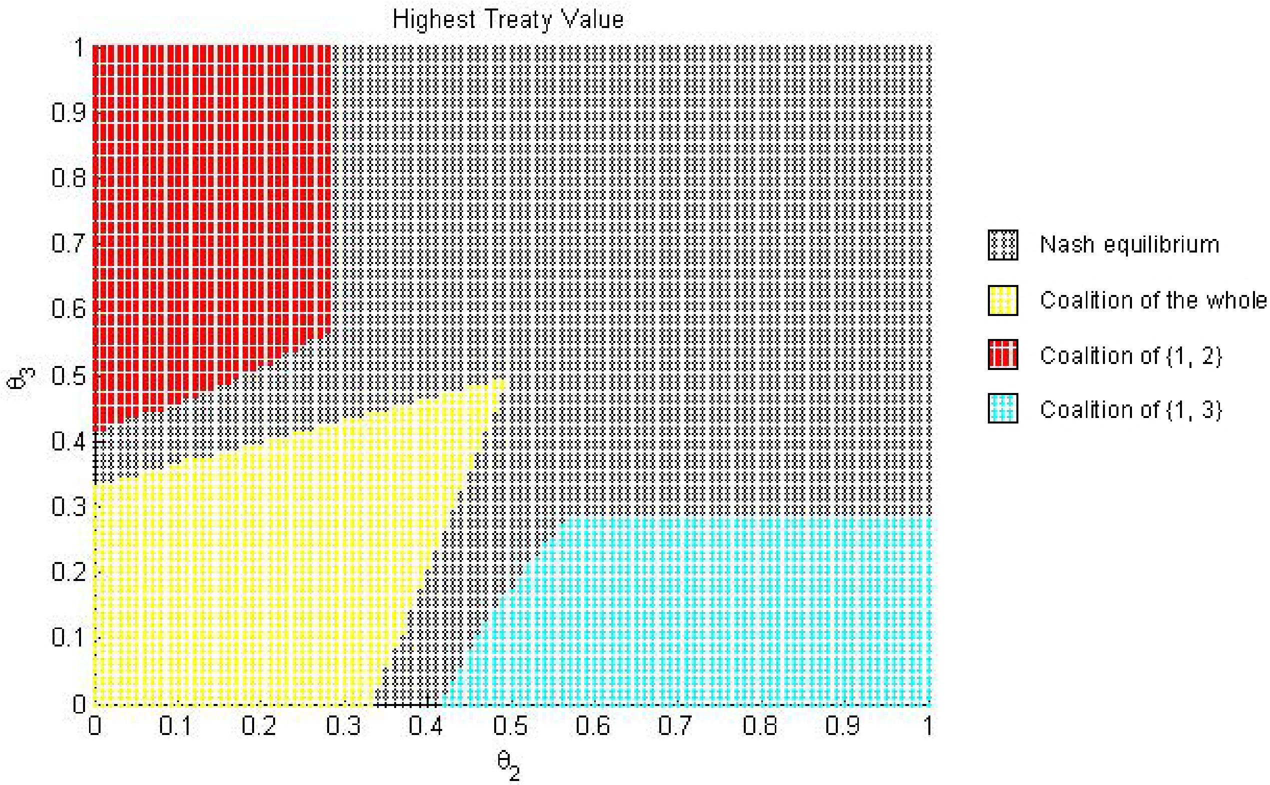

FIGURE 1. θ_1 is normalized to zero.

where player 1 takes much larger action than player 2 or 3, has no treaties which improve upon the Nash equilibrium for all players.

This example gives a clear view of how more homogeneous agents can band together to improve total utility. When all three players are similar, they form the coalition of the whole; when one agent is less similar, he is excluded from treaty negotiation. This result is generalized in the next section.

4. Model and Analysis

This section describes a set of negative externality games played by coalitions with different restrictions on their actions. I first specify the constituent elements, then the classes of games.

4.1. General Set-Up.

I study equilibria of games in which coalitions have commitment power of different sorts. In all of the games, there is a set of agents I, with cardinality n at least equal to three. Each agent's action set is $A_i = [0, 1]$, with $A \equiv \times_{i \in I} A_i$, and the utility functions belong to the class $\mathcal{U}$ satisfying the following conditions:

a. twice continuously differentiable, each u_i is in $C^2(A)$,
b. negative externalities, $(\forall i \in I)(\forall j \neq i)(\forall a \in A)\left[\frac{\partial u_i(a)}{\partial a_j} < 0\right]$,
c. strict submodularity, $(\forall i \in I)(\forall j \neq i)(\forall a \in A)\left[\frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} < 0\right]$,
d. strict own concavity, $(\forall i \in I)\left[\frac{\partial^2 u_i(a)}{\partial a_i^2} < 0\right]$, and
e. strong dominant effect, $(\forall i \in I)(\forall a \in A)\left[\left|\frac{\partial^2 u_i(a)}{\partial a_i^2}\right| > \left|\sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j}\right|\right]$,
f. unique interior Nash equilibrium, the Kuhn-Tucker conditions for equilibrium of u have a unique solution on the interior of A.

These elements describe a fairly general class of negative externality games. In a game without any cooperation, the Nash equilibria involve only the singleton coalitions.

Definition 1. **[Nash Equilibrium.]** *For utility function $u \in \mathcal{U}$, a^* is a* ***Nash equilibrium at*** u*, denoted* $a^*(u) \in Eq(u)$*, if for all* $i \in I$*,* $a_i^*(u)$ *solves* $\max_{b_i \in A_i} u_i(b_i, a_{-i}^*)$.

The negative externalities condition guarantees that any Nash equilibrium is inefficient and reductions strictly improve everyone's welfare.

Lemma 1. *For any* $u \in \mathcal{U}$*, if* $a^*(u) \in Eq(u)$*, then any small vector decrease in* $a^*(u)$ *is Pareto improving.*

This result relies on demonstrating that each agent's small decrease in action has a first order effect on others' utility but only a second order effect on own utility. Generalizations of this idea can be found in Anderson and Zame [3].

4.2. **Games Played by Coalitions.** Lemma 1 establishes that a group of agents may form to act together. Group formation is typically modeled via coalition games. Multilateral treaties are agreements to cooperate in the interests of group welfare, so I will examine games played by coalitions to gain insights on treaty formation. I ignore the details of the bargaining process in favor of finding agreements that make everyone party to them better off.

Prior to taking action, agents are invited to negotiate a single agreement; there are no side agreements or alternate provisions possible. Coalitions can vary in commitment power, and I study two types of commitment: first, agents in a coalition can agree to a specific vector of commitments which lists the action taken by each member of the final agreement; second, agents in a coalition can agree to a one-dimensional reduction from the Nash equilibrium. However, the externality in this class of games gives rise to a free-rider problem. If a coalition J forms and commits to reductions, the players not in the coalition will increase their outputs in response, because of the higher marginal utility resulting from the strict submodularity of the utility function and the decreased actions of the coalition members.

Joining a coalition must give some benefit to the participants. Therefore, coalitions are certainly not even conceivable if they do not do as well for as the Nash equilibrium their members; such coalitions simply would not form.

Definition 2. **[Conceivability.]** *A vector of actions* $a(u, J)$ *is* **conceivable** *for coalition* J *if each agent within* J *experiences a weak improvement in utility over the Nash equilibrium and at least one agent experiences a strict improvement. Formally,*

$$(\forall j \in J)\left[u_j\left(a(u,J)\right) \geq u_j\left(a^*(u)\right)\right] \tag{5}$$

and

$$(\exists k \in J)\left[u_k\left(a(u,J)\right) > u_k\left(a^*(u)\right)\right]. \tag{6}$$

At issue is how much reduction will be achieved by various coalitions when they have different kinds of commitment power. I begin by studying unlimited commitment power, then turn to the ability to commit to a lump-sum reduction from the Nash equilibrium.

4.2.1. *Unrestricted Coalition Commitment.* Unrestricted coalition play games for a coalition $\emptyset \neq J \subseteq I$ are the games played by the agents in J acting as a single player with summed group utility function $\sum_{j \in J} u_j$ and every $i \notin J$ acting as a single player with utility function u_i.

Definition 3. **[Unrestricted Coalition Power.]** *For* $J \subseteq I$*, the* ***J-coalition game with unrestricted coalition commitment power****,* $\Gamma_J^{Un}(u)$*, has* $\#(I \setminus J) + 1$ *agents, with agent* J *having action set* $\times_{j \in J} A_j$ *with utility function* $u_J(a) = \sum_{j \in J} u_j(a)$*, and agents* $i \notin J$ *having action sets* A_i *with utility functions* $u_i(a)$*. The equilibrium of this game is denoted* $a^{Un}(u, J) = (a_J^{Un}(u), a_{-J}^{Un}(u))$.[1]

As earlier discussed, Lemma 1 establishes that a reduction from Nash by all players will improve utility for each player. A coalition of the whole could most easily achieve such an outcome in a game with unrestricted commitment power. By maximizing group utility, such a coalition is not only conceivable but efficient as well. No other coalition in unrestricted commitment power can improve upon it.

[1]Note that if $J = \{i\}$ is a coalition of size 1, then $a^{Un}(u, J)$ is equal to $a^*(u)$, the Nash equilibrium.

Lemma 2. *For $u \in \mathcal{U}$, when coalitional commitment power is unrestricted, the unrestricted equilibrium of the coalition of the whole, $a^{Un}(u, I)$, is conceivable for all $u \in \mathcal{U}$, and no other coalition J strictly smaller than I can improve upon upon the actions in summed utility.*

The best possible coalition under unrestricted actions is the coalition of the whole, though coalitions excluding players with Nash action of zero may tie in utility if the excluded players still have a best response of zero.

4.2.2. *Restricted Coalition Commitment.* Apart from the unrestricted ability of assigning an individual target to each agent, the vector of coalition commitments can be constructed in a few manners. One possibility is to establish a one-dimensional decrease from previous actions, for instance each by some equal lump-sum reduction or by some equal percentage reduction. Requiring all members to follow the same reduction rule has a sense of egalitarianism and is often observed in real world agreements, like the proportional reduction in the Montreal Protocol [6]. In this paper, I define lump-sum commitment power.[2]

Consider the game where a coalition can commit to lump-sum reductions. I suppose a Nash equilibrium, $a^*(u)$ prevails, and a coalition J commits to a lump-sum reduction of r. By this I mean that each agent in J reduces their Nash action by an amount r.

Definition 4. [Lump-sum Coalition Power.] *For $J \subseteq I$, the **J-coaltion game with lump-sum commitment power**, $\Gamma_J^{LS}(u)$, has $\#(I \setminus J) + 1$ agents, with agent J having the action set $\{a_J \in \times_{j \in J} A_j : (\forall j \in J)[a_j = a_j^*(u) - r], r \in [0, \min_j a_j^*(u)]\}$ with utility function $u_J(a) = \sum_{j \in J} u_j(a)$, and agents $i \notin J$ having action sets A_i with utility functions $u_i(a)$. The equilibrium of this game is denoted $a^{LS}(u, J) = (a_J^{LS}(u), a_{-J}^{LS}(u))$.*

This solution concept is subtle because the equilibrium definitions have Nash equilibria within them. This is why uniqueness of the Nash equilibrium is so important. The solution concept could be weakened to non-unique Nash games, perhaps by choosing the "largest" equilibrium, or by ignoring relabeled equilibria. Despite this being a static game, the negotiations could be thought of as if the players are agreeing to a per-period action, with the Nash equilibrium as a fall-back.

Lemma 2 illustrated that the coalition of the whole is the best possible option in terms of group utility in the case of unrestricted commitment power. However, this coalition may be thwarted if the gains to free-riding are especially high, in the presence of uncertainty, or in a dynamic game without sufficient patience. Furthermore, under the restriction of lump-sum commitment power, the coalition of the whole may not be ideal. Sufficient heterogeneity in the costs and benefits of agents guarantees that, with lump-sum commitment power, coalitions strictly smaller than I are conceivable and and also improve upon a coalition of the whole. One can go even further, for some vectors of payoff functions $(u_i)_{i \in I}$, and find the set(s) J that are Pareto superior to the coalition of the whole amongst all subsets of I.

The following result demonstrates that there are coalitions J, strictly smaller than the full set of agents, with restricted commitment power which improve upon the Nash equilibrium and upon the result of the whole coalition. Together, these give the result that coalitions strictly smaller than I are conceivable and Pareto-improving when there is enough heterogeneity and when actions are restricted in this reasonable manner.

Theorem 1. *For any $J \subsetneq I$, $\#J \geq 2$, there is a set of $u \in \mathcal{U}$ having non-empty interior, for which the vector of actions $a^{LS}(u, J)$ is conceivable, formally denoted as:*

$$(\forall j \in J)[u_j(a^{LS}(u, J)) > u_j(a^*(u))]. \tag{7}$$

Further, there is a subset of $u \in \mathcal{U}$ having non-empty interior which fulfill the above and for which, under the lump-sum restriction, the coalition J improves upon the outcome of the coalition of the whole, formally written as:

$$(\forall i \in I)[u_i(a^{LS}(u, J)) > u_i(a^{LS}(u, I))]. \tag{8}$$

[2]Other possible restrictions include a proportional commitment, where agents each decrease by some percentage, or an upper limit $\bar{X}$ on the stock of negative actions and then split it according to some sharing rule. A carbon cap program contains stock limits, though such a program is not so much a treaty as an implemented policy. In a smaller example, home owner's association members must restrict all noise to a lower decibel level at nighttime.

The proof can be found in the Appendix. The next section elaborates on the intuition of this result using a class of models with parametrized utility functions.

4.3. **Parametrized Class of Utility Functions.** To clarify the exclusion result given in Theorem 1, I discuss a utility function similar to those common in the literature on coalitions as environmental agreements. The functional form can be separated into a benefit function and a damage function.

As initially presented in Section 4.1, the game has agents $I = \{1, 2, ..., n\}$, where each takes action $a_i \in A_i = [0, 1]$. These players are heterogeneous in the following fashion: each player $i \in I$ has a positive benefits coefficient, $\theta_i \in \Theta_i = [0, 1)$, which multiplies the benefit gained from the action taken. Thus, the class of utility functions examined here consists of those which have the following form:

$$u_i(a_i, a_{-i}) = (1 - \theta_i)B(a_i) - a_i\, c\left(\sum_{k\in I} a_k\right)$$

Here, $B(a_i)$ represents the benefits of the individual action, with multiplicative coefficient $(1 - \theta_i)$. The function is increasing and concave, i.e. $B' > 0, B'' < 0$. The cost of individual action is $a_j\, c\left(\sum_k a_k\right)$, where the marginal cost depends on the weighted summed total action.[3] The marginal cost function is increasing and convex, i.e. $c' > 0, c'' > 0$.

The benefit and damage functions are shared among players, and together they must fulfill the characteristics defined earlier on $\mathcal{U}_i$, which were negative externalities, strict submodularity, strict own concavity, strong dominant effect, and unique interior Nash equilibrium. The benefit and cost functions presented can easily fulfill all of these requirements together, and each of the characteristics can be checked when functional forms and number of agents are assigned.

There are two groups of players:

1. The first is group J with cardinality m. The agents in group J all have $\theta_i = 0$. Hence, the utility function for players $j \in J$ is:
$$u_j(a_j, a_{-j}) = B(a_j) - a_j c\left(\sum_{k\in I} w_{jk} a_k\right) \tag{9}$$
2. The second group of players $I \setminus J$ consists of the remaining $(n - m)$ players. These agents all have $\theta_i = \theta$. The utility function for a player $i \in I \setminus J$ is:
$$u_i(a_i, a_{-i}) = (1 - \theta)B(a_i) - a_i c\left(\sum_{k\in I} w_{ik} a_k\right) \tag{10}$$

The parameter θ decreases the marginal benefit of the action as it increases from zero to one. As θ increases, group heterogeneity grows larger, and the two groups grow more disparate. In equilibrium, the players in J should be taking the same action, as should all the players in $I \setminus J$.

4.3.1. *Nash Equilibrium.* Suppressing the u from the notation in the previous sections, the Nash equilibrium of this game $a^*(\theta)$ consists of the equilibrium actions of the players in $I \setminus J$, denoted $a^*_{I\setminus J}(\theta)$, and the equilibrium actions of the players in J, denoted $a^*_J(\theta)$:

$$a^*_{I\setminus J}(\theta) \equiv \arg\max_{a_i\in A_i}\ (1 - \theta)B(a_i) - a_i c\big(a_i + (n - m - 1)\, a^*_{I\setminus J}(\theta) + m a^*_J(\theta)\big), \tag{11}$$

$$a^*_J(\theta) \equiv \arg\max_{a_j\in A_j} B(a_j) - a_j c\big(a_j + (n - m)a^*_{I\setminus J}(\theta) + (m - 1)a^*_J(\theta)\big) \tag{12}$$

At the lowest value of the group parameter, $\theta = 0$, both groups of players have the same maximization problem. Since the players are identical in this case, they would play the same action: $a^*_{I\setminus J}(0) = a^*_J(0)$. At higher values of the group parameter, $\theta > 0$, the two groups of agents have different maximization problems and different actions.

Lemma 3. *When the group parameter is strictly positive, $\theta > 0$, then the equilibrium action of the players in $I \setminus J$, $a^*_{I\setminus J}(\theta)$, is smaller than the equilibrium action of the players in J, $a^*_J(\theta)$.*

[3]Often, the action vector could be left unsummed in this function or, in a setting of incomplete information, this sum is possibly a discounted expectation [12].

The proof shows that a strictly positive θ decreases the marginal benefit of action of the players in $I \setminus J$ compared to that of the players in J.

According to Lemma 1 from Section 4.1, this Nash equilibrium is inefficient. Therefore, the agents may form a coalition in which they agree to reduce the action and total negative externality. Under a coalition with unrestricted power, the agents could easily achieve a first-best solution, assigning a specified action to each agent. Under a coalition with lump-sum restricted power, each agent participating must subtract the same amount from his Nash action. Examining restricted power coalitions in this class of utility functions will help illustrate the exclusion result in Theorem 1.

4.3.2. *Lump-sum Reductions.* The agents consider two possible lump-sum reduction treaties: one which forms a coalition of the whole (all I players), and one which contains only the players in J (excluding those in $I \setminus J$)

First, consider the coalition of all I players. The coalition maximization problem is:

$$\max_{r \in \left[0, a^*_{I\setminus J}(\theta)\right]} \Big\{ \textstyle\sum_{i \in I \setminus J} (1-\theta) B\big(a^*_{I\setminus J}(\theta) - r\big) - \big(a^*_{I\setminus J}(\theta) - r\big) c \Big((n-m)\big(a^*_{I\setminus J}(\theta) - r\big) + m\big(a^*_J(\theta) - r\big) \Big)$$
$$+ \textstyle\sum_{j \in J} B\big(a^*_J(\theta) - r\big) - \big(a^*_J(\theta) - r\big) c \Big((n-m)\big(a^*_{I\setminus J}(\theta) - r\big) + m\big(a^*_J(\theta) - r\big) \Big) \Big\} \tag{13}$$

For every θ and equilibrium $a^*(\theta)$, there is some utility-maximizing lump-sum reduction for the coalition of the whole denoted $r^*_I(\theta)$. This reduction solves the first order condition listed in the Appendix. The choice of $r^*_I(\theta)$ is limited by the smaller action, $a^*_{I\setminus J}(\theta)$. The action space is bounded from below by 0, so the coalition's reduction can only be as large as the smallest action of a participant, meaning that $r^*_I(\theta) \le a^*_{I\setminus J}(\theta)$.

For small values of θ, this condition does not pose a problem. If the optimal reduction for the coalition of the whole is strictly smaller than the Nash action chosen by the players in $I \setminus J$, then the optimal reduction is implemented and all players participate. However, consider the case as θ approaches one. Then, the equilibrium action of the players in $I \setminus J$ approaches $a^*_{I\setminus J}(1) \to 0$, which limits the reduction that a coalition of the whole could implement.

The agents in J still take a strictly positive action. Though the agents in $I \setminus J$ have negligible actions, the players in J continue to exert a negative externality on each other and could agree to reduce by themselves. A separate treaty for the players in J would benefit all players.

Thus, consider the coalition of only the players in J. For any value of θ, the coalition maximization problem is:

$$\max_{r \in \left[0, a^*_J(\theta)\right]} \ \textstyle\sum_{j \in J} B\big(a^*_J(\theta) - r\big) - \big(a^*_J(\theta) - r\big) c \Big((n-m) a^J_{I\setminus J}(\theta) + m\big(a^*_J(\theta) - r\big) \Big) \tag{14}$$

while agents in $I \setminus J$ best respond as singletons with:

$$a^J_{I\setminus J}(\theta) \equiv \arg\max_{a_i \in A_i} (1-\theta) B(a_i) - a_i\, c \Big(a_i + (n-m-1) a^J_{I\setminus J}(\theta) + m\big(a^*_J(\theta) - r^*_J(\theta)\big) \Big) \tag{15}$$

For every θ and equilibrium $a^*(\theta)$, there is some utility-maximizing lump-sum reduction for the coalition of the whole denoted $r^*_J(\theta)$. This reduction solves the first order condition listed in the Appendix.

At any θ, the J-coalition's marginal utility to increasing reduction from zero is strictly positive, i.e. $\frac{\partial u_J}{\partial r}|_{r=0} > 0$ (shown in Appendix A.2), given that the players in $I \setminus J$ were best responding by playing Nash. Even if the free-riding players increase actions from Nash equilibrium by a tiny amount, the J-coalition will still be Pareto improving for the agents in J. This hints at the fact that, when the agents in $I \setminus J$ have minimal response, the J coalition has $r^*_J(\theta)$ strictly greater than zero and that the exclusion result holds for high values of θ.

In the coalition of whole, total utility is increasing in the reduction r for some time, and then begins to decrease. The first order condition for the I-coalition gives weight to the marginal benefit to action of each group according to size of that group, choosing a reduction between what would be optimal for those in $I \setminus J$ and those in J. At $\theta = 1$, the marginal benefit of action is zero for agents in $I \setminus J$, so attempting to include them restricts the possibilities of reduction. Moving away from this high θ, the marginal utility of those in $I \setminus J$ turns negative more quickly than the marginal utility of those in J, preserving the exclusion result.

Lemma 4. *There exists a threshold value $\underline{\theta} < 1$ for the group parameter such that for all values of the parameter higher than the threshold, $\theta \in (\underline{\theta}, 1)$, the equilibrium action of the players in $I \setminus J$, $a^*_{I \setminus J}(\theta)$, is a binding constraint on maximization problem (13).*

Lemma 4 demonstrates that the exclusion result from Theorem 1 holds for this class of parametrized utility functions. Furthermore, it elucidates the mechanics of the exclusion result: for zero actions, players are excluded because they simply cannot take smaller actions; for larger actions, players are excluded because they *will not* take smaller actions, since their marginal utility would run negative. Either way, the smallest actions bind the egalitarian action space of any coalition which would include them.

4.3.3. *Multi-level Coalitions.* What if the coalition of the whole optimized among multiple levels of one-dimensional action? Agents may be open to the possibility of multi-level coalitions, meaning that all participants belong to the same coalition but that different actions are prescribed for segmented levels of actors. For instance, large action actors may be given one lump-sum reduction, while small action actors are given another. In the case of this parametrized class, where much of the heterogeneity stems from a group parameter, a multi-level coalition could be a way to include the players in $I \setminus J$ in a coalition of the whole even when θ is close to one.

However, with only two types of players, specifying two levels of reduction is akin to having unrestricted commitment power. One can see how multi-level coalitions converge to unrestricted commitment power. As more and more levels are introduced, the coalition can tailor action to each member, in the same way that unrestricted actions would be constructed. This eliminates the very object of interest, which is the response of players to a single, egalitarian action. Furthermore, in a game with uncertainty, a multi-level coalition may still end up having only one level, if agents wish to obscure type.

5. Conclusions

Heterogeneity plays a great role in the MP constraint chosen. If agents are more homogeneous, then an increase in the size of the MP constraint will benefit the treaty. With more heterogeneous agents in the world, agents with large actions follow the intuition of excluding the "odd man out" and creating a Pareto-improving treaty, as opposed to signing an inclusive but less effective treaty. In a fairly general class of negative externality games, there are groups of agents strictly smaller than the coalition of the whole which perform strictly better under a lump-sum reduction constraint. This translates to an MP constraint strictly smaller than n, rendering small-action agents non-pivotal. In an application to environmental treaties, this could translate to large polluters using MP constraints to exclude smaller polluters and form a better performing MEA. This result is also evidenced by the use of smaller negotiation spaces, which for instance would not take place at the UN, but at exclusive summits.

Such an egalitarian restriction creates an interesting paradox where agents most damaged by the externality or with least benefit from taking the action are most eager to limit the total stock but cannot join an egalitarian lump-sum reduction treaty. Furthermore, as can be seen throughout the examples of the paper, there is some residual inefficiency to this egalitarian approach. In the motivating example of Section 3, unable to include a third agent in the treaty when his production is low, the other two must play Nash equilibrium in order to keep him from free-riding on the public good by increasing in response. A coalition with unrestricted commitment power could ask the third agent to simply stay at zero level, and the treaty would form. This example illustrates the very tension between Lemma 2 and Theorem 1 – one specifies that a strictly smaller than whole coalition cannot improve upon everyone, while the other result gives exactly the opposite. The restricted actions drive the exclusion result, which in many cases will give a coalition arbitrarily close to efficiency.

In this model, agents are indifferent between ineffective treaties and no treaty at all, allowing a symbolic agreement to be equilibrium behavior. To avoid symbolic equilibria, these could be discouraged with a minor tie-breaking rule or minimal cost to entering negotiations, or encouraged with some sort of utility boost from the appearance of concern. However, it is rather intriguing that those with the most to gain from effective treaties are the ones that can join only symbolic treaties when actions are restricted, since they have the least to contribute.

The fact remains that greater treaty membership does not automatically mean greater treaty benefit. This is clearly illustrated in overfishing control: the promise of landlocked countries to limit their fishing has little meaning when their access to the fishing stock is already limited by geography.

Of great hope is the result that positive producers prefer effective treaties to symbolic ones, at least in certain regions of improvement. The MP mechanism is somewhat successful in internalizing the common damages faced by agents and attracts participants by increasing the benefit of the treaty itself, even in a one-shot game. Nevertheless, the increase in initial value offered by the MP constraint can be combined with other self-enforcement mechanisms to enhance not just treaty participation, but treaty effectiveness as well.

References

[1] Europe.eu: Money and the EU, January 2013.

[2] Europe.eu: The history of the European Union, January 2013.

[3] Robert M. Anderson and William R. Zame. Genericity with infinitely many parameters. *Advances in Theoretical Economics*, 1(1):1, 2001.

[4] Kyle Bagwell. Self-enforcing trade agreements and private information. Technical report, National Bureau of Economic Research, 2009.

[5] Scott Barrett. Self-enforcing international environmental agreements. *Oxford Economic Papers*, 46:878–894, 1994.

[6] Kurt J. Beron, James C. Murdoch, and Wim P. M. Vijverberg. Why cooperate? Public goods, economic power, and the Montreal Protocol. *Review of Economics and Statistics*, 85(2):286–297, 2003.

[7] Jane Black, Maurice D. Levi, and David de Meza. Creating a good atmosphere: Minimum participation for tackling the 'Greenhouse Effect'. *Economica*, 60(239):281–293, 1993.

[8] Carlo Carraro, Carmen Marchiori, and Sonia Oreffice. Endogenous minimum participation in international environmental treaties. *Environmental and Resource Economics*, 42(3):411–425, 2009.

[9] Ward Cheney. Calculus in Banach spaces. In *Analysis for Applied Mathematics*, volume 208 of *Graduate Texts in Mathematics*, pages 115–169. 2001.

[10] Dean Corbae, Maxwell B. Stinchcombe, and Juraj Zeman. *An introduction to mathematical analysis for economic theory and econometrics*. Princeton University Press, 2009.

[11] Peter Egger, Christoph Jeßberger, and Mario Larch. The determinants of multilateral environmental agreement membership. Technical report, Discussion paper, mimeo.

[12] Étienne Billette de Villemeur and Justin Leroux. Looking ahead versus looking back: Revisiting the carbon tax!, 2011.

[13] Alexander Hamilton. Federalist paper no. 1. 1787.

[14] Anna A. Klis. Early participants of multilateral environmental agreements and their effect on later membership. In *Carroll Round Proceedings*, volume 6, pages 238–251. Georgetown University, 2010.

[15] Roger Lagunoff and Rodrigo Harrison. Dynamic mechanism design for a global commons. *Available at SSRN*, 2013.

[16] Rodney D. Ludema and Anna Maria Mayda. Do countries free ride on MFN? *Journal of International Economics*, 77(2):137–150, 2009.

[17] Ulrich J. Wagner. The design of stable international environmental agreements: Economic theory and political economy. *Journal of Economic Surveys*, 15(3):377–411, 2001.

[18] Hans-Peter Weikard, Leo Wangler, and Andreas Freytag. Minimum participation rules with heterogeneous countries. Technical report, Jena economic research papers, 2009.

Appendix A. Proofs and Equations

A.1. **Expanded Results from Section 4.1.**

Proof of Lemma 1.

Restatement of Lemma 1. *For any $u \in \mathcal{U}$, if $a^*(u) \in Eq(u)$, then any small vector decrease in $a^*(u)$ is Pareto improving.*

Proof. For each i, evaluated at $a^*(u)$, an agent i's marginal utility of his own action $\frac{\partial u_i(\cdot)}{\partial a_i} = 0$, so an ε-decrease from a_i^* will reduce i's utility by something on the order of ε^2. However, for all agents $j \neq i$, the marginal utility for j of i's action is strictly negative $\frac{\partial u_j(\cdot)}{\partial a_i} < 0$, so the ε-decrease from a_i^* will increase j's utility on the order of ε. For small ε, $\varepsilon^2 < \varepsilon$. □

Proof of Lemma 2.

Restatement of Lemma 2. *For $u \in \mathcal{U}$, when coalitional commitment power is unrestricted, the unrestricted equilibrium of the coalition of the whole, $a^U n(u, I)$, is conceivable for all $u \in \mathcal{U}$, and no other coalition J strictly smaller than I can improve upon upon the actions in summed utility.*

Proof. By Lemma 1, $a^{Un}(u, I)$ is conceivable. By definition, the unrestricted coalitional commitment power solves the problem, $\max_{a \in A} \sum_{i \in I} u_i(a)$, where each member can be assigned any action in the space, so no other vector of actions, equilibrium or not, gives a higher sum. □

Proof of Theorem 1.

Restatement of Theorem 1. *For any $J \subsetneq I$, $\#J \geq 2$, there is a set of $u \in \mathcal{U}$ having non-empty interior, for which the vector of actions $a^{LS}(u, J)$ is conceivable, formally denoted as:*

$$(\forall j \in J)[u_j(a^{LS}(u, J)) > u_j(a^*(u))].$$

Further, there is a subset of $u \in \mathcal{U}$ having non-empty interior which fulfill the above and for which, under the lump-sum restriction, the coalition J improves upon the outcome of the coalition of the whole, formally written as:

$$(\forall i \in I)[u_i(a^{LS}(u, J)) > u_i(a^{LS}(u, I))].$$

Proof. Pick some J with cardinality greater than one. Pick a function u fulfilling all of the desired characteristics and where exclusion is optimal under the lump-sum restriction. This proof will use a particular function from that class for which all the assumptions on $\mathcal{U}$ hold, and then demonstrate that those assumptions are open conditions.

The particular function considered is

$$u_i(a) = (1 - \theta_i)(10 + a_i) - a_i^2 \left(n + \sum_{j=1}^{n} a_j \right). \tag{16}$$

The parameter θ is as described in Section 4.3, the group J consists of agents who have $\theta_i = 0$, while the remaining agents in $I \setminus J$ have $\theta_i = \theta \in \Theta = (0, 1)$.

Exclusion Result

See the proof of Lemma 4.

Fulfillment of Assumptions

- *Twice continuous differentiability:* This assumption clearly holds for this utility function, as the first and second total and partial derivatives can be easily taken.

– First Derivatives

$$\frac{du_i(a)}{da} = \frac{\partial u_i(a)}{\partial a_i} + \sum_{j \neq i} \frac{\partial u_i(a)}{\partial a_j}$$
$$\frac{\partial u_i(a)}{\partial a_i} = (1-\theta_i) - \left[2a_i\left(n + \sum_{j=1}^{n} a_j\right) + a_i^2\right]$$
$$\forall\, j \neq i \;\; \frac{\partial u_i(a)}{\partial a_j} = -a_i^2$$

– Second Derivatives

$$\frac{d^2 u_i(a)}{da^2} = \frac{\partial^2 u_i(a)}{\partial a_i^2} + 2\sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} + \sum_{j \neq i} \left\{ \frac{\partial^2 u_i(a)}{\partial a_j^2} + \sum_{k \neq i \text{ or } j} \frac{\partial u_i(a)}{\partial a_j \partial a_k} \right\}$$
$$\frac{\partial^2 u_i(a)}{\partial a_i^2} = -\left[2\left(n + \sum_{j=1}^{n} a_j\right) + 4a_i\right]$$

$$\forall\, j \neq i \;\; \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} = -2a_i$$
$$\forall\, j \neq i \;\; \frac{\partial u_i(a)}{\partial a_j^2} = 0$$
$$\forall\, k \neq i \text{ or } j \;\; \frac{\partial u_i(a)}{\partial a_j \partial a_k} = 0$$

- *Negative externalities:* Using the derivatives above, I can confirm negative externalities on the domain. The first derivative of i's utility function with respect to any j's action is:

$$\forall\, j \neq i \;\; \frac{\partial u_i(a)}{\partial a_j} = -a_i^2$$

For every action in the set $A_i = [0,1]$, this derivative is less than or equal to zero. It is strictly negative for actions in $(0,1]$ and only zero when no action is taken, i.e. $a_i = 0$.

- *Submodularity:* Again, using the derivatives above, I can confirm submodularity on the the domain. The cross-partial of i's utility function with respect to his own action and another agent j's action is:

$$\forall\, j \neq i \;\; \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} = -2a_i$$

For every action in the set $A_i = [0,1]$, this derivative is less than or equal to zero. It is strictly negative for actions in $(0,1]$ and only zero when no action is taken, i.e. $a_i = 0$.

- *Strict own concavity:* Once again, using the derivatives above, I can confirm strict concavity on the the domain. The second derivative of i's utility function with respect to his own action is:

$$\frac{\partial^2 u_i(a)}{\partial a_i^2} = -\left[2\left(n + \sum_{j=1}^{n} a_j\right) + 4a_i\right]$$

For any $a \in A$, this derivative is strictly negative.

- *Strong dominant effect:* To confirm strong dominant effect, compare the second derivative of u_i with respect to i with the sum of the cross partials with respect to $j \neq i$.

$$\left|\frac{\partial^2 u_i(a)}{\partial a_i^2}\right| \overset{?}{>} \left|\sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j}\right|$$
$$\left|-\left[2\left(n + \sum_{j=1}^{n} a_j\right) + 4a_i\right]\right| \overset{?}{>} \left|\sum_{j \neq i} -2a_i\right|$$
$$\left[2\left(n + \sum_{j=1}^{n} a_j\right) + 4a_i\right] \overset{?}{>} 2(n-1)a_i$$
$$2n + 2\sum_{j=1}^{n} a_j + 4a_i \overset{?}{>} 2(n-1)a_i$$

$n > n-1$ and $a_i \leq 1 \Rightarrow 2n > 2(n-1)a_i$, and $2\sum_{j=1}^{n} a_j + 4a_i \geq 0 \Rightarrow$

$$2n + 2\sum_{j=1}^{n} a_j + 4a_i > 2(n-1)a_i$$

Hence, strong dominant effect holds for all i and a.

- *Unique interior Nash equilibrium:* The Nash maximization problem for an agent, given other's Nash actions a_j^*, is the following:

$$\begin{aligned} \max_{a_i \in [0,1]} \; & (1-\theta_i)(10+a_i) - a_i^2 \left(n + \sum_{j=1}^{n} a_j \right) \\ \text{s.t. } & a_i \geq 0 \\ & a_i \leq 1 \end{aligned}$$

The Lagrangian is:

$$\mathcal{L}(a_i, \lambda_{1i}, \lambda_{2i}) = (1-\theta_i)(10+a_i) - a_i^2 \left(n + \sum_{j=1}^{n} a_j \right) + \lambda_{1i} a_i + \lambda_{2i}(1-a_i)$$

The Kuhn-Tucker conditions are:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial a_i} =& (1-\theta_i) - \left[2a_i \left(n + \sum_{j=1}^{n} a_j \right) + a_i^2 \right] + \lambda_{1i} - \lambda_{2i} = 0 \\ & \lambda_{1i} a_i = 0, \quad \lambda_{1i} \geq 0, a_i \geq 0 \\ & \lambda_{2i}(1-a_i) = 0, \quad \lambda_{2i} \geq 0, a_i \leq 1 \end{aligned}$$

There are three cases to examine: interior solution, corner solution of zero, and corner solution of one.

Case i. Interior action: $a_i \in (0,1) \Rightarrow \lambda_{1i} = \lambda_{2i} = 0$
The first derivative of the Lagrangian becomes:

$$\begin{aligned} \left[2a_i \left(n + \sum_{j=1}^{n} a_j \right) + a_i^2 \right] &= (1-\theta_i) \\ \left[2a_i \left(n + \sum_{j \neq i} a_j^* \right) + 3a_i^2 \right] &= (1-\theta_i) \\ 3a_i^2 + 2a_i \left(n + \sum_{j \neq i} a_j^* \right) - (1-\theta_i) &= 0 \end{aligned}$$

Using the quadratic formula to solve for the optimal action:

$$\begin{aligned} a_i^* &= \frac{-2\left(n + \sum_{j \neq i} a_j^* \right) \pm \sqrt{4\left(n + \sum_{j \neq i} a_j^* \right)^2 + 12(1-\theta_i)}}{6} \\ &= \frac{-\left(n + \sum_{j \neq i} a_j^* \right) \pm \sqrt{\left(n + \sum_{j \neq i} a_j^* \right)^2 + 3(1-\theta_i)}}{3} \end{aligned}$$

Check if these answers are interior. First, check a_i^{*-} (which uses the minus from $\pm$):

$$a_i^{*-} = \frac{1}{3}\left[-\left(n + \sum_{j \neq i} a_j^* \right) - \sqrt{\left(n + \sum_{j \neq i} a_j^* \right)^2 + 3(1-\theta_i)} \right]$$

$$\begin{aligned} \text{Know } & \left(n + \sum_{j \neq i} a_j^* \right)^2 + 3(1-\theta_i) > 0 \\ \Rightarrow \; & \sqrt{\left(n + \sum_{j \neq i} a_j^* \right)^2 + 3(1-\theta_i)} > 0 \\ \Rightarrow \; & \frac{1}{3}\left[-\left(n + \sum_{j \neq i} a_j^* \right) - \sqrt{\left(n + \sum_{j \neq i} a_j^* \right)^2 + 3(1-\theta_i)} \right] < 0 \end{aligned}$$

This means that a_i^{*-} is not within the domain, and so it cannot be a Nash equilibrium action. Now check a_i^{*+} (which uses the plus from $\pm$):

$$a_i^{*+} = \frac{1}{3}\left[-\left(n+\sum_{j\neq i} a_j^*\right)+\sqrt{\left(n+\sum_{j\neq i} a_j^*\right)^2+3(1-\theta_i)}\right]$$

$$\text{Know} \quad \sqrt{\left(n+\sum_{j\neq i} a_j^*\right)^2+3(1-\theta_i)} > \left(n+\sum_{j\neq i} a_j^*\right)$$

$$\Rightarrow \quad \frac{1}{3}\left[-\left(n+\sum_{j\neq i} a_j^*\right)+\sqrt{\left(n+\sum_{j\neq i} a_j^*\right)^2+3(1-\theta_i)}\right] > 0$$

This means that a_i^{*+} is greater than zero; now check if it is less than one.

$$\frac{1}{3}-\left(n+\sum_{j\neq i} a_j^*\right)+\sqrt{\left(n+\sum_{j\neq i} a_j^*\right)^2+3(1-\theta_i)} \overset{?}{<} 1$$

$$-\left(n+\sum_{j\neq i} a_j^*\right)+\sqrt{\left(n+\sum_{j\neq i} a_j^*\right)^2+3(1-\theta_i)} \overset{?}{<} 3$$

$$\sqrt{\left(n+\sum_{j\neq i} a_j^*\right)^2+3(1-\theta_i)} \overset{?}{<} \left(n+\sum_{j\neq i} a_j^*\right)+3$$

$$\left(n+\sum_{j\neq i} a_j^*\right)^2+3(1-\theta_i) \overset{?}{<} \left(\left(n+\sum_{j\neq i} a_j^*\right)+3\right)^2$$

$$\left(n+\sum_{j\neq i} a_j^*\right)^2+3(1-\theta_i) \overset{?}{<} \left(n+\sum_{j\neq i} a_j^*\right)^2+6\left(n+\sum_{j\neq i} a_j^*\right)+9$$

$$3(1-\theta_i) < 6\left(n+\sum_{j\neq i} a_j^*\right)+9$$

Since $(1-\theta_i)<1$, this certainly holds. Thus, we know that $a_i^{*+}<1$.

Case ii. Corner solution of zero: $a_i^*=0 \Rightarrow \lambda_{2i}=0$

The first derivative of the Lagrangian becomes:

$$(1-\theta_i)+\lambda_{1i}=0$$
$$\Rightarrow (1-\theta_i)\leq 0$$
$$\Rightarrow \theta_i = 1$$

This is outside the range of $\Theta_i=[0,1)$, so this case will not occur.

Case iii. Corner solution of one: $a_i^*=1 \Rightarrow \lambda_{1i}=0$

The first derivative of the Lagrangian becomes:

$$(1-\theta_i)-\left[2\left(n+\sum_{j\neq i} a_j\right)+3\right]-\lambda_{2i}=0$$

$$(1-\theta_i)=\left[2\left(n+\sum_{j\neq i} a_j\right)+3\right]+\lambda_{2i}$$

$$\Rightarrow (1-\theta_i)\geq 2\left(n+\sum_{j\neq i} a_j\right)+3$$

However, since $1-\theta_i\leq 1$ and $2\left(1+\sum_{j\neq i} a_j^*\right) >> 1$, this case can never occur.

Thus the interior case is the only one which will be chosen. Now I show that the equilibrium is unique through proof by contradiction. Suppose there exists a^* and $a^{*}*$ s.t. that the interior Kuhn-Tucker conditions are fulfilled, i.e. for all i both of the following hold:

$$2a_i^*\left(n+\sum_{j=1}^n a_j^*\right)+a_i^{*2}=(1-\theta_i)$$

$$2a_i^{**}\left(n+\sum_{j=1}^n a_j^{**}\right)+a_i^{**2}=(1-\theta_i)$$

Summing these conditions over i, the following two conditions must hold:

$$2\left(n+\sum_{i=1}^n a_i^*\right)\sum_{i=1}^n a_i^* + \sum_{i=1}^n a_i^{*2} = \sum_{i=1}^n (1-\theta_i)$$
$$2\left(n+\sum_{i=1}^n a_i^{**}\right)\sum_{i=1}^n a_i^{**} + \sum_{i=1}^n a_i^{**2} = \sum_{i=1}^n (1-\theta_i)$$

Subtract the bottom condition from the top one:

$$\left[2\left(n+\sum_{i=1}^n a_i^*\right)\sum_{i=1}^n a_i^* + \sum_{i=1}^n a_i^{*2}\right] - \left[2\left(n+\sum_{i=1}^n a_i^{**}\right)\sum_{i=1}^n a_i^{**} - \sum_{i=1}^n a_i^{**2}\right] = 0$$

$$2\left[\left(n+\sum_{i=1}^n a_i^*\right)\sum_{i=1}^n a_i^* - \left(n+\sum_{i=1}^n a_i^{**}\right)\sum_{i=1}^n a_i^{**}\right] + \left[\sum_{i=1}^n a_i^{*2} - \sum_{i=1}^n a_i^{**2}\right] = 0$$

$$2n\left[\sum_{i=1}^n a_i^* - \sum_{i=1}^n a_i^{**}\right] + 2\left[\left(\sum_{i=1}^n a_i^*\right)^2 - \left(\sum_{i=1}^n a_i^{**}\right)^2\right] + \left[\sum_{i=1}^n a_i^{*2} - \sum_{i=1}^n a_i^{**2}\right] = 0 \qquad (17)$$

This can only be solved if $\sum_{i=1}^n a_i^* = \sum_{i=1}^n a_i^{**}$ and $\sum_{i=1}^n a_i^{*2} = \sum_{i=1}^n a_i^{**2}$. However, this condition does not yet imply that the two equilibria are equal, i.e. that $a_i^* = a_i^{**}$ for all i.

In order for a^* and a^{**} to **not** be the same, there must be at least one person for whom the actions are different. Without loss of generality, suppose $a_i^* \neq a_i^{**}$. Check the conditions for i to see whether this is possible.

$$2a_i^*\left(n+\sum_{j=1}^n a_j^*\right) + a_i^{*2} = (1-\theta_i)$$

$$2a_i^{**}\left(n+\sum_{j=1}^n a_j^{**}\right) + a_i^{**2} = (1-\theta_i)$$

Recall that the total agent sums must be the same, i.e. $\sum_{i=1}^n a_i^* = \sum_{i=1}^n a_i^{**}$. Subtract the bottom condition from the top one:

$$2(a_i^* - a_i^{**})\left(n+\sum_{j=1}^n a_j^*\right) + (a_i^{*2} - a_i^{**2}) = 0$$

Substitute the factorization for the sum of squares:

$$2(a_i^* - a_i^{**})\left(n+\sum_{j=1}^n a_j^*\right) + (a_i^* + a_i^{**})(a_i^* - a_i^{**}) = 0$$

If $a_i^* \neq a_i^{**}$, this means we can divide through by $(a_i^* - a_i^{**})$, since it is not equal to zero. This gives:

$$2\left(n+\sum_{j=1}^n a_j^*\right) + (a_i^* + a_i^{**}) = 0$$

$$2\left(n+\sum_{j=1}^n a_j^*\right) = -(a_i^* + a_i^{**})$$

This lead to a contradiction: $n > 0$ and for all j, $a_j^* \geq 0$, meaning that the left-hand side is strictly positive, while the right-hand side must be weakly negative. Therefore it must be that $a_i^* = a_i^{**}$ for all i, meaning that a^* and a^{**} are the same and that the equilibrium is unique.

Openness of Conditions

The C^2-norm on the utility functions for $i \in I$ is:

$$||u_i||_i \equiv \max_{a\in A}|u_i(a)| + \sum_j \max_{a\in A}\left|\frac{\partial u_i(a)}{\partial a_j}\right| + \sum_{k,j\in I,k\geq j}\max_{a\in A}\left|\frac{\partial^2 u_i(a)}{\partial a_k \partial a_j}\right|$$

The product space of the individual utility function is $u \in C^2(A;\mathbb{R}^I)$, with the norm:

$$||u|| \equiv \max_{i\in I}||u_i||_i$$

The distance between two utility outcomes is $d(u,v) \equiv ||u-v||$. To show openness, I will show that for a $u \in \mathcal{U}$ there exists $\varepsilon > 0$ such that $||u-v|| < \varepsilon$ implies that $v \in \mathcal{U}$ also.

- *Negative Externality*

  The problem $\max_{i,j}\max_a \frac{\partial u_i(a)}{\partial a_j}$ has a solution $\{i^*, j^*, a^*\}$. Since $u \in \mathcal{U}$, $\frac{\partial u_{i^*}(a^*)}{\partial a_{j^*}} < 0$. Define

  $$\varepsilon_a \equiv \frac{1}{2}\left|\frac{\partial u_{i^*}(a^*)}{\partial a_{j^*}}\right|.$$

  If $||u-v|| < \varepsilon_a$, then $\forall\, i, \forall\, j, \forall\, a$ $\frac{\partial v_i(a)}{\partial a_j} < -\varepsilon_a < 0$. Thus, v has negative externalities.

- *Submodularity and Concavity*

  Similarly as above, the problem $\max_{i,j} \max_a \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j}$ has a solution $\{i^*, j^*, a^*\}$. Since $u \in \mathcal{U}$, $\frac{\partial^2 u_{i^*}(a^*)}{\partial a_{i^*} \partial a_{j^*}} < 0$. Define

$$\varepsilon_b \equiv \frac{1}{2} \left| \frac{\partial^2 u_{i^*}(a^*)}{\partial a_{i^*} \partial a_{j^*}} \right|.$$

  If $||u - v|| < \varepsilon_b$, then $\forall\, i, \forall\, j \in I$ (including i), $\forall\, a$ $\frac{\partial^2 v_i(a)}{\partial a_i \partial a_j} < -\varepsilon_b < 0$. Thus, v is submodular and concave.

- *Strong Dominant Effect*

Lemma A1. *Suppose $u \in \mathcal{U}$ and v is concave and submodular for all i, j, a. Then $||u - v|| < \varepsilon \Rightarrow v$ also has the strong dominant effect property.*

We know that u has $\forall i$, $\forall j$, $\forall a$ $\left|\frac{\partial^2 u_i(a)}{\partial a_i^2}\right| > \left|\sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j}\right|$. The problem $\max_i \max_a \left(\frac{\partial^2 u_i(a)}{\partial a_i^2} - \sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j}\right)$ has a solution $\{i^*, a^*\}$. Since both quantities are negative, and the own second derivative is larger in absolute value (therefore more negative), this maximum value is negative. Define

$$\varepsilon_c \equiv \frac{1}{2} \left(\left| \frac{\partial^2 u_{i^*}(a^*)}{\partial a_{i^*}^2} \right| - \left| \sum_{j \neq i^*} \frac{\partial^2 u_{i^*}(a^*)}{\partial a_{i^*} \partial a_j} \right| \right).$$

If $||u - v|| < \varepsilon_c$, then $\forall\, i, \forall\, j, \forall\, a$:

$$\left| \frac{\partial^2 u_i(a)}{\partial a_i^2} - \frac{\partial^2 v_i(a)}{\partial a_i^2} \right| + \left| \sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} - \sum_{j \neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j} \right| < \varepsilon_c$$

By definition $\left|\frac{\partial^2 u_i(a)}{\partial a_i^2}\right| - \left|\sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j}\right| \geq 2\varepsilon_c$. Add and subtract $\left|\frac{\partial^2 v_i(a)}{\partial a_i^2}\right|$:

$$\left| \frac{\partial^2 u_i(a)}{\partial a_i^2} \right| + \left| \frac{\partial^2 v_i(a)}{\partial a_i^2} \right| - \left| \frac{\partial^2 v_i(a)}{\partial a_i^2} \right| - \left| \sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} \right| \geq 2\varepsilon_c$$

Since these are all negative values, they can be combined within the absolute value signs:

$$\left| \frac{\partial^2 u_i(a)}{\partial a_i^2} - \frac{\partial^2 v_i(a)}{\partial a_i^2} \right| + \left| \frac{\partial^2 v_i(a)}{\partial a_i^2} \right| - \left| \sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} \right| \geq 2\varepsilon_c$$

Now add and subtract $\sum_{j \neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j}$:

$$\left| \frac{\partial^2 u_i(a)}{\partial a_i^2} - \frac{\partial^2 v_i(a)}{\partial a_i^2} \right| + \left| \frac{\partial^2 v_i(a)}{\partial a_i^2} \right| - \left| \sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} \right| + \left| \sum_{j \neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j} \right| - \left| \sum_{j \neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j} \right| \geq 2\varepsilon_c$$

Again, these can be recombined within the absolute value signs:

$$\left| \tfrac{\partial^2 u_i(a)}{\partial a_i^2} - \tfrac{\partial^2 v_i(a)}{\partial a_i^2} \right| + \left| \tfrac{\partial^2 v_i(a)}{\partial a_i^2} - \textstyle\sum_{j \neq i} \tfrac{\partial^2 v_i(a)}{\partial a_i \partial a_j} \right| - \left| \textstyle\sum_{j \neq i} \tfrac{\partial^2 u_i(a)}{\partial a_i \partial a_j} \right| + \left| \textstyle\sum_{j \neq i} \tfrac{\partial^2 v_i(a)}{\partial a_i \partial a_j} \right| \geq 2\varepsilon_c$$

$$\left| \tfrac{\partial^2 u_i(a)}{\partial a_i^2} - \tfrac{\partial^2 v_i(a)}{\partial a_i^2} \right| + \left| \tfrac{\partial^2 v_i(a)}{\partial a_i^2} - \textstyle\sum_{j \neq i} \tfrac{\partial^2 v_i(a)}{\partial a_i \partial a_j} \right| + \left| \textstyle\sum_{j \neq i} \tfrac{\partial^2 v_i(a)}{\partial a_i \partial a_j} \right| - \left| \textstyle\sum_{j \neq i} \tfrac{\partial^2 u_i(a)}{\partial a_i \partial a_j} \right| \geq 2\varepsilon_c$$

$$\left| \tfrac{\partial^2 u_i(a)}{\partial a_i^2} - \tfrac{\partial^2 v_i(a)}{\partial a_i^2} \right| + \left| \tfrac{\partial^2 v_i(a)}{\partial a_i^2} - \textstyle\sum_{j \neq i} \tfrac{\partial^2 v_i(a)}{\partial a_i \partial a_j} \right| + \left| \textstyle\sum_{j \neq i} \tfrac{\partial^2 v_i(a)}{\partial a_i \partial a_j} - \textstyle\sum_{j \neq i} \tfrac{\partial^2 u_i(a)}{\partial a_i \partial a_j} \right| \geq 2\varepsilon_c$$

$$\left| \tfrac{\partial^2 u_i(a)}{\partial a_i^2} - \tfrac{\partial^2 v_i(a)}{\partial a_i^2} \right| + \left| \tfrac{\partial^2 v_i(a)}{\partial a_i^2} - \textstyle\sum_{j \neq i} \tfrac{\partial^2 v_i(a)}{\partial a_i \partial a_j} \right| + \left| \textstyle\sum_{j \neq i} \tfrac{\partial^2 u_i(a)}{\partial a_i \partial a_j} - \textstyle\sum_{j \neq i} \tfrac{\partial^2 v_i(a)}{\partial a_i \partial a_j} \right| \geq 2\varepsilon_c$$

Using the earlier fact that $\left|\frac{\partial^2 u_i(a)}{\partial a_i^2} - \frac{\partial^2 v_i(a)}{\partial a_i^2}\right| + \left|\sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} - \sum_{j \neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j}\right| < \varepsilon_c$, we can see that:

$$\left|\frac{\partial^2 v_i(a)}{\partial a_i^2} - \sum_{j \neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j}\right| \geq 2\varepsilon_c - \left|\frac{\partial^2 u_i(a)}{\partial a_i^2} - \frac{\partial^2 v_i(a)}{\partial a_i^2}\right| - \left|\sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} - \sum_{j \neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j}\right|$$

$$\left|\frac{\partial^2 v_i(a)}{\partial a_i^2} - \sum_{j \neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j}\right| > 2\varepsilon_c - \varepsilon_c$$

$$\left|\frac{\partial^2 v_i(a)}{\partial a_i^2} - \sum_{j \neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j}\right| > \varepsilon_c$$

Therefore $\forall\, i, \forall\, j, \forall\, a$ $\left|\frac{\partial^2 v_i(a)}{\partial a_i^2}\right| - \left|\sum_{j \neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j}\right| > 0$, and strong dominant effect holds for v.

- *Unique Interior Nash Equilibrium*

Lemma A2. *Suppose $u \in \mathcal{U}$ and v is concave and submodular and has negative externalities and strong dominant effect for all i, j, a. Then there exists ε such that $||u - v|| < \varepsilon \Rightarrow v$ also has a unique Nash equilibrium.*

Proceed with proof by contradiction. Let $\mathcal{U}'$ be the set of C^2 functions that are concave, submodular, and have negative externalities and strong dominant effect, but may have multiple interior Nash equilibria. This new set is a superset of $\mathcal{U}$. Suppose that $v \in \mathcal{U}'$ has more than one equilibrium. I will show that if it is close to u, then this means that u should also have multiple equilibria.

I use the **General Implicit Function Theorem** (Theorem 3) from Ward [9]: *Let X, Y, and Z be normed linear spaces, Y being assumed complete. Let Ω be an open set in $X \times Y$. Let $F : \Omega \to Z$. Let $(x_0, y_0) \in \Omega$. Assume that F is continuous at (x_0, y_0), that $F(x_0, y_0) = 0$, that D_2F exists in Ω, that D_2F is continuous at (x_0, y_0), and that $D_2F(x_0, y_0)$ is invertible. Then there is a function f defined on a neighborhood of x_0 such that $F(x, f(x)) = 0$, $f(x_0) = y_0$, f is continuous at x_0, and f is unique in the sense that any other such functions must agree with f on some neighborhood of x_0.*

Denote the following:

- $X = \mathcal{U}'$
- $Y = [0,1]^n$
- $Z = \mathbb{R}^n$
- $\Omega \subseteq (0,1)^n$
- $F = \begin{pmatrix} \frac{\partial v_1(a)}{\partial a_1} \\ \frac{\partial v_2(a)}{\partial a_2} \\ \vdots \\ \frac{\partial v_n(a)}{\partial a_n} \end{pmatrix}$
- (x_0, y_0) are the interior Nash equilibria $(v, a^*(v))$ and $(v, a^{**}(v))$

Since F represents the interior Kuhn-Tucker conditions (or first order conditions), we have that:

$$F(v, a^*(v)) = \begin{pmatrix} \frac{\partial v_1(a^*(v))}{\partial a_1} \\ \frac{\partial v_2(a^*(v))}{\partial a_2} \\ \vdots \\ \frac{\partial v_n(a^*(v))}{\partial a_n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Each first derivative is continuous, so F is continuous at all points. Since $\mathcal{U}' \in C^2$, D_2F exists and can be written as:

$$D_2F = \begin{pmatrix} \frac{\partial^2 v_1(a)}{\partial a_1^2} & \frac{\partial^2 v_1(a)}{\partial a_1 \partial a_2} & \cdots & \frac{\partial^2 v_1(a)}{\partial a_1 \partial a_n} \\ \frac{\partial^2 v_2(a)}{\partial a_1 \partial a_2} & \frac{\partial^2 v_2(a)}{\partial a_2^2} & \cdots & \frac{\partial^2 v_2(a)}{\partial a_2 \partial a_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 v_n(a)}{\partial a_1 \partial a_n} & \frac{\partial^2 v_n(a)}{\partial a_2 \partial a_n} & \cdots & \frac{\partial^2 v_n(a)}{\partial a_n^2} \end{pmatrix}$$

This matrix exists throughout $[0,1]^n$ and is continuous at all points. Finally, because of the property of strong dominant effect, D_2F is a diagonally dominant matrix, ensuring that it is invertible throughout the domain.

The conditions for the the General Implicit Function Theorem are fulfilled. Therefore, there is some function f which assigns Nash equilibria on a neighborhood of v which are close to $a^*(v)$, and some function f' which assigns equilibra which are close to $a^{**}(v)$. Thus, for $u \in \mathcal{U}$ which is also within some ε_e-neighborhood of v, there must be $a^*(u)$ and $a^{**}(u)$ which are respectively close to $a^*(v)$ and $a^{**}(v)$ which are also Nash equilibria, i.e. satisfy $F(u, a^*(u)) = F(u, a^{**}(u)) = 0$. This is a contradiction that u has a unique interior Nash equilibrium. Hence, v has only one interior equilibrium.

Each of the conditions has been shown to be open individually. Take $\varepsilon^* \equiv \min\{\varepsilon_a, \varepsilon_b, \varepsilon_c, \varepsilon_d\}$. Pick v such that $||u - v|| < \varepsilon^*$. Then all the conditions are satisfied by v. Hence, the set of utility functions in $\mathcal{U}$ is open. This means the exclusion result holds on an open set.

□

A.2. **Expanded Results from Section 4.2.**

First Order Conditions for Coalitions choosing a lump-sum reduction:

- Coalition of the whole, I

$$\begin{aligned}\frac{n-m}{n}&(1-\theta)B'\left(a^*_{I\setminus J}(\theta) - r^*_I(\theta)\right) + \frac{m}{n}B'\left(a^*_J(\theta) - r^*_I(\theta)\right) = \\ &\left[(m-n)\left(a^*_{I\setminus J}(\theta) - r^*_I(\theta)\right) - m\left(a^*_J(\theta) - r^*_I(\theta)\right)\right] c'\left((m-n)\left(a^*_{I\setminus J}(\theta) - r^*_I(\theta)\right) + m\left(a^*_J(\theta) - r^*_I(\theta)\right)\right) \\ &+ c\left((m-n)\left(a^*_{I\setminus J}(\theta) - r^*_I(\theta)\right) + m\left(a^*_J(\theta) - r^*_I(\theta)\right)\right)\end{aligned}$$

- Coalition of agents in J

$$\begin{aligned}B'\left(a^*_J(\theta) - r^*_J(\theta)\right) =& c\left((m-n)a^J_{I\setminus J}(\theta) + m(a^*_J(\theta) - r^*_J(\theta))\right) \\ &+ m\left(a^*_J(\theta) - r^*_J(\theta)\right) c'\left((m-n)a^J_{I\setminus J}(\theta) + m(a^*_J(\theta) - r^*_J(\theta))\right)\end{aligned}$$

Incentive for J-coalition to Reduce.

Lemma A3. *When moving from zero, the players in J in the game in Section 4.3 have a strict incentive to increase r and reduce. Formally, this means that the following statement is positive:*

$$\left.\frac{\partial u_J}{\partial r}\right|_{r=0} = -m\left[B'\left(a^*_J(\theta)\right) - c\left((m-n)a^J_{I\setminus J}(\theta) + m a^*_J(\theta)\right) - m a^*_J(\theta) c'\left((m-n)a^J_{I\setminus J}(\theta) + m a^*_J(\theta)\right)\right] \quad (18)$$

Proof. From the Nash first order conditions, $a^*_J(\theta)$ solves:

$$B'\left(a^*_J(\theta)\right) - c\left((m-n)a^*_{I\setminus J}(\theta) + m a^*_J(\theta)\right) - a^*_J(\theta) c'\left((m-n)a^*_{I\setminus J}(\theta) + m a^*_J(\theta)\right) = 0$$

At $r^*_I = 0$, those in J are agreeing to play Nash, meaning that the best responses of those in $I \setminus J$ are also Nash, giving $a^*_{I\setminus J}(\theta) = a^J_{I\setminus J}(\theta)$. The Nash first order condition and the derivative of u with respect to

r evaluated at zero are then nearly identical, apart from the extra weight on the cost derivative. Comparing the two, then it must be that:

$$B'\left(a_J^*(\theta)\right) - c\left((m-n)a_{I\setminus J}^J(\theta) + ma_J^*(\theta)\right) - ma_J^*(\theta)c'\left((m-n)a_{I\setminus J}^J(\theta) + ma_J^*(\theta)\right) < 0$$

Combined with the negative sign on the outside of the parentheses, Statement (18) is positive. □

<u>Proof of Lemma 4.</u>

Restatement of Lemma 4. *There exists a threshold value $\underline{\theta} < 1$ for the group parameter such that for all values of the parameter higher than the threshold, $\theta \in (\underline{\theta}, 1]$, the equilibrium action of players in $I \setminus J$, $a_{I\setminus J}^*(\theta)$, is a binding constraint on problem 13.*

Proof. I show that there exists an open set around 1 in the parameter range Θ for which $r_J^*(\theta)$ is strictly positive and large compared to $r_I^*(\theta)$, which is close to zero.

1. *Continuity:* The first step is to show that the Nash equilibrium $a^*(\theta)$ is continuous. I use the result that if and only if some function $\Phi : X \to \mathcal{K}(Y)$ has a closed graph and Y is compact, then Φ is upper hemi-continuous (Corollary 6.1.33 in [10]).

   Here, Φ is our equilibrium correspondence, defined below as f. The X is the game Γ, which is defined below. The Y is the paramter $\theta \in \Theta = [0, 1]$ (which is immediately observed to be compact), and $\mathcal{K}(Y)$ is the set of strategy profiles $\sigma \in \Delta(A)$.

   1.1 *Establishing closed graph:*
   - Each player has a utility function, $u_i : A \times \Theta \to \mathbb{R}$.
   - The game is defined as follows: $\Gamma(\theta) \equiv \{(u_i(\cdot;\theta), A_i)_{i\in I} : \theta \in \Theta\}$.
   - A set of strategies, σ^* is a Nash equilibrium of the game Γ if σ_i^* performs at least as well as any other strategy a_i^o for player i given that the other agents are playing σ_{-i}^*. Formally, $\sigma^* \in \mathrm{Eq}(\Gamma(\theta))$ if and only if for all agents i in all possible sets of agents I and for all alternate strategies $a_i^o \in A_i$, then:

   $$f(\sigma, \theta; a_i^o) \equiv \int_A u_i(a;\theta)d\sigma^*(a) - \int_A u_i(a \setminus a_i^o)d\sigma^*(a) \geq 0 \tag{19}$$

   Then the graph is defined $\mathrm{Gr}(f) \equiv \{(\sigma, \theta) : f(\sigma, \theta; a_i^o) \geq 0\}$, and it is closed.

   Thus, the equilibrium correspondence of the game $\Gamma(\theta)$ is upper hemi-continuous. Furthermore, a function that is upper hemi-continuous and single-valued at a point is continuous at that point. Since the equilibrium correspondence is upper hemi-continuous, and the optimal actions $a_J^*(\theta)$ and $a_{I\setminus J}^*(\theta)$ of Equations (9) and (10) are single-valued for each θ, then the equilibrium correspondence is single-valued for each θ. Thus, the equilibrium correspondence of the game is continuous.

2. *Close to zero:* The next step is to assert that equilibria for values of the parameter strictly inside the parameter space may be close to zero. By continuity, since at the boundary parameter value of one and the equilibrium action $a_{I\setminus J}^*(1) = 0$, then $a_{I\setminus J}^*(\theta)$ for θ arbitarily close to one is arbitrarily close to zero.

3. *Binding:* The third step is to show that when solving the coalition I problem for θ close to 1, $a_{I\setminus J}^*(\theta)$ is a binding constraint on choosing $r_I^*(\theta)$. Looking at Equation 13, it balances the utility of both groups of players. The players in $I \setminus J$ can only decrease to zero, meaning that the entire problem is constrained by the size of $a_{I\setminus J}^*(\theta)$. However, the players in J have a strictly positive benefit from group reduction, as shown in Proposition A3. Thus, the action of the players in $I \setminus J$ is a binding constraint on the coalition of the whole's reduction problem.

Since the reduction that can be implemented by the coalition of the whole is constrained to be very small because the actions of the agents in $I \setminus J$ is very small, the players in J will prefer to form the J coalition (positive incentive to reduction, very low free-riding by non-members). This improves the utility of those in J *and* those in $I \setminus J$, giving a Pareto improvement upon the coalition of the whole for some values of the parameter.

□