The Dissertation Committee for Anna Alexandra Klis
certifies that this is the approved version of the following dissertation:

Essays on Externalities and International Cooperation: A Game Theoretic Approach

Committee:

Maxwell B. Stinchcombe, Supervisor

Thomas E. Wiseman

Dale O. Stahl

Stephen P. Ryan

Scott J. Moser

Essays on Externalities and International Cooperation: A Game Theoretic Approach

by

Anna Alexandra Klis, B.S. For. Ser., M.S. Econ.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2015

Dla moich rodziców, za ich cierpliwość, miłość, i trzymanie kciuków.

To my parents, for their patience, love, and finger-crossing.*

*Literally translated, “held thumbs” – idiomatically, however, for wishing me the best of luck and keeping me in their thoughts.

Acknowledgments

First and foremost, I must thank Maxwell Stinchcombe for his teaching and advising. Without his curiosity or his support, I would have been unable to tackle these theoretical problems that "bite back."

Second, I would like to thank Thomas Wiseman for his guidance in improving my job market paper and presentation and for his help in translating my ideas for a general audience. I would also like to thank Dale Stahl and participants in UT's theory writing seminars, the 25th Stony Brook International Conference on Game Theory, and AEA/ASSA 2015 for their comments aiding the development of these papers, as well as Stephen Ryan and Scott Moser for providing extra perspective.

Finally, I would like to thank Ina Taneva, Melinda Petre, Matthew Shapiro, Virginia Luerhsen, and my parents, for their valuable assistance in the editing process; Esther Raizen, Associate Dean in the UT College of Liberal Arts, and summer 2014 dissertation boot camp participants for their inspiration and insights; and Amanda Delp, for her inestimable help with the final stages of this project.

Essays on Externalities and International Cooperation: A Game Theoretic Approach

Publication No. ____________________

Anna Alexandra Klis, Ph.D.
The University of Texas at Austin, 2015

Supervisor: Maxwell B. Stinchcombe

In this dissertation, I present three essays which examine questions in the field of public economics using a game theoretic approach, and I derive hopeful results and helpful rules for international negotiation.

In my first chapter, I examine minimum participation constraints. In the presence of heterogeneity, a minimum participation (MP) clause in a public goods arrangement can serve as a device to create a more homogeneous group. When coalitions are restricted in what they can bargain over, exclusion of some agents from the bargaining process can be Pareto improving. This paper gives a general set of sufficient conditions for such an exclusion result to hold, and presents examples of when exclusion does, and does not, improve upon unanimity.

In the second chapter, I discuss the problem of determining which externality situations merit international cooperation. I create a general framework

of linearized parameters to examine a general externality problem, and then I provide the sufficient conditions for a parameter to move non-cooperative and cooperative solutions in opposite directions under certain circumstances. I argue that situations which behave in this manner and which have a higher parameter value have more benefit to cooperation through the increased range in actions to bargain over.

The third chapter extends upon the second chapter and applies the framework developed to an externality problem. I present a particular story of correlation in fish growth and a corresponding model which gives an example of an increasing action gap. I describe the method of use of the framework, and using the linearized parameters developed in the second chapter, I attempt to show the divergence of non-cooperative and cooperative actions in this setting, demonstrating the need for negotiation among sovereign entities.

Table of Contents

List of Figures

Chapter 1

Minimum Participation Clauses and Exclusion in Public Good Agreements

1.1 Introduction

An agreement for the abatement of a public bad or the supply of a public good requires determination of the exact action to be taken and the set of agents to be involved. Internationally, a typical multilateral environmental agreement (MEA) may be negotiated to reduce transboundary pollution, increase fishing stocks, or control regional radioactive waste. A Home Owners' Association, on the other hand, has a goal to reduce annoying actions amongst the community and may involve a single apartment complex or a whole block of houses. Though the agreements operate on different levels of agents, the goal to reduce a negative externality is the same.[1] Other examples of coalitions reducing a negative externality are the European Union, which originally had the aim of ending "frequent and bloody wars between neighbors," and the

[1] It is true that agents most often encounter previously established Home Owners' Associations, while just-forming Associations usually have a very clear set of agents to draw from when a new development is built. Agents joining MEAs, on the other hand, are usually directly involved in determining the provisions during negotiation. However, the parallel stands, because there are Home Owners' Associations which encompass residential homes, as opposed to buildings, and the lines of membership must be drawn, and there is accession to existing treaties.

euro currency zone, which was partly intended to limit the impact of currency exchange on trade deals and tourist experiences [4, 5].

Sovereignty implies that agents cannot be forced to do something by another party's will alone. In an environment with sovereign agents, lack of external enforcement means that only mutually beneficial agreements can correct an externality. Countries are fully sovereign, since a country cannot force its will on another without making war and since there is no ruling third party. Citizens do not have quite the same degree of sovereignty, since they need to follow a government's laws, but laws rarely extend to cosmetic issues, such as curtain lining color, or establish less stringent guidelines than a community would impose, as in the case of noise controls.[2] For either country or citizen, there must be a significant benefit to joining a coalition and binding oneself 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 [37].

If sovereign agents take a hand in designing their own agreement, then agents would endogenously determine the actions taken, as well as the self-enforcement mechanisms implemented. First, agents must choose the method and amount of contribution to the public good, and they must do so in a way that is in every participant's best interests. Different situations may specify for

[2]Another difference to note is that in a Home Owner's Association, individuals fully commit, while in a multilateral environmental agreement, citizens and firms may violate an alloted share of some sort. This paper will therefore treat countries as the individual, and assume perfect enforcement for them upon their commitment.

agents varying ability to commit to certain types of group action. In the Home Owners' Association, agents can commit to a monthly dues system easily, but keeping communal areas clean may be more difficult. In the MEA, agents may agree on the necessity of cooperation and sharing of research, but may be hard-pressed to give specific details of policy.

Second, the mechanisms which will make the agreement stick and increase its value must be determined. Some possible provisions that enhance participation in agreements are minimum participation clauses, direct transfers, and issue linkages. 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 of nine 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.

In a public-good contribution game, the optimal binding treaty would specify an action for each agent, even if the agents differ in their contribution costs. If the set of possible contracts is limited, though – principally, if contracts must specify the same increase in contributions for all agents – then the optimal contract may exclude agents with a high cost of contributing. This is significant because important examples of such asymmetry are present in overfishing and nuclear armament: in the former, countries are limited in types of fishing by geography, while costs may vary widely because of technology; in the latter, countries have different historical starting positions, needs for nuclear power, and perceived intentions. Therefore, though all agents may have interest in the overarching topic and benefit from the public good, restrictions in what the agents bargain over affects the optimal MP constraint in the presence of heterogeneous players.

This paper examines how restricted action sets affect the MP constraint that should be chosen for a treaty in the presence of heterogeneous players. In particular, it evaluates how agreement actions which are restricted to egalitarianism, either in the form of equal changes or proportional changes from the pre-treaty state, 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 treaties with higher total welfare.

In Section 1.2, 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 1.3, I present a motivating example, while in Section 1.4, I set up my general model, define two types of restricted action set, present results, and describe the intuition behind them. In Section 1.5, I compare the two types of restricted action sets and place them in a real-world context. In Section 1.6, I present conclusions and possible extensions. Finally, I have included an Appendix that contains the major and minor proofs, as well as one with alternative proof approaches.

1.2 Literature

Barrett [10] presents two approaches to modeling 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 a treaty relying on these strategies 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 [15] examine the introduction of an exogenous MP constraint into a one-shot game. They find that inclusion of any constraint larger than the resulting number of members under open participation outperforms a standard agreement. In their model, the new constraint increases the number of participants and lowers the total stock of the negative action. The constraint can be constructed so as to maximize aggregate surplus, resulting in an optimal participation level. 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, since pivotal signers wish to gain further benefits.

Carraro, Marchiori, and Oreffice [20] 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 from the coalition formed under open membership. Here, already, there is some notion of exclusion; an MP constraint requiring the coalition of the whole may not be chosen due to the incentive to be a free-rider. Because each agent wants a chance to free-ride, he supports an MP constraint which gives him some chance of joining the agreement and some chance to strictly benefit. However, the analysis is sensitive to the assumption of homogeneity of agents: the MP stage can be solved with unanimous voting since what is optimal in the eyes of one agent is optimal for all of them.

While arguments using models with identical agents capture the important aspects of many situations, in others, heterogeneity of costs and benefits is of central importance in the analysis of the actions to which signatories will bind themselves. Agents can vary in terms of 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 [13, 28, 43]. A source of heterogeneity among countries for the issue of pollution is that of technology, since a country on the cutting edge of technology likely has lower costs of reduction compared to a country with little research and development. Meanwhile, in the case of nuclear disarmament, the largest source of heterogeneity is preference for security and perceived threat level. On a smaller level, when establishing an Home Owners' Association, families may value different restrictions, such as noise control or cleanliness of public areas, than do single households. A Home Owners' Association formed among similar family-size homes may find it easier to enact certain restrictions than a community of apartments of varied occupancy.[3]

Weikard, Wangler, and Freytag [59] extend Carraro et al.'s model to heterogeneous agents. They use the same coalition formation timing as Carraro et al., but change the minimum participation constraint from number of signatories to minimum abatement, which has some precedence. Weikard et al. designate a sharing rule proportional to outside options to determine the actions of any coalition. A random agent is chosen to propose which agents

[3]This, of course, is a moot point if the Association's rules are established by a developer prior to construction as part of a marketing plan to select for a homogeneous group of homeowners.

should enter a coalition under this sharing rule, capturing the idea that someone's proposal will win, but it is hard to predict whose. In this set-up, they find that free-riding always occurs, at least by one agent – the agenda setter. In addition, they find that a larger number of countries leads to a smaller abatement outcome which is inefficient.

Using a coalition formation model to represent a treaty negotiation process separates the MP and allocation decisions. The MP is chosen in a first stage through the statement of a minimal coalition, ahead of the allocation decision, which is set in a following stage by the formed coalition. This may be the proper timing for certain applications such as the home owners' association or even the euro zone, when membership is established before rules. However, in most multilateral international negotiations, countries write agreements over a period of time and then vote on all final provisions in one shot. Initially, one might suppose that using incorrect timing may limit the theoretical predictions and outside relevance of the model. Despite that, 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 the equilibria of a one-shot negative externality game and the set of agreements that improve upon the no-coalition Nash equilibrium. This can be understood as an equilibrium in a repeated game with Nash reversion, though I do not develop that idea or pursue the enforcement of agreements. In the MEA context, the one-shot game would be a single treaty negotiation, which must be adhered to or no other one-shot games can be played in the nebulous future. Thus, it is as if agents were embedded in a larger game of international politics and respect for negotiations, where cheating on the outcome of a one-shot treaty coalition game leads to collapse of the system.

Like Weikard et al. [59], I consider a specific type of treaty action. Unlike their sharing rule, I develop a solution concept under the limited commitment power of an equal treatment assumption, where coalition members can only commit to one-dimensional decreases from the *ex ante* no-coalition Nash equilibrium. Under this exogenous constraint of egalitarianism, I examine the MP choices that give the most improvement over no-coalition 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 can contribute a bit more of the public good and improve the utility of all participants. Even though it adds a restriction, egalitarianism can be a desirable treaty trait. Requiring all coalition members to take the same action allows for simplicity in negotiation, since the choice variable can be one-dimensional instead of multi-dimensional.

Furthermore, egalitarianism may result from environments with uncertainty. In a dynamic externality reduction game with private cost shocks, Harrison and Lagunoff [44] find that truth-telling and coalition participation require fully compressed quotas, i.e. amount allowances which cannot depend on private information, but must the same for all players. Agents are initially identical, even if later they develop heterogeneously. Regardless of later shocks, all agents in the agreement have the same per-period production quota as other agents. Bagwell [9] develops a bilateral tariff negotiation game with uncertainty in types, solving 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. The results of both of these papers add to the motivation of understanding the use of egalitarian treaties.

The equal treatment assumption changes the structure and participation of an enacted treaty in comparison to unrestricted actions. The main result of this paper is that under egalitarianism and given sufficient heterogeneity, the optimal MP constraint is strictly smaller than the 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 [48] find a similar exclusion result in their paper on tariff negotiations within the World Trade Organization. The WTO's most-favored nation status must apply to all members. If a large exporter with most-favored nation status negotiates a lower tariff with an importer, then the lower tariff applies to all members and results in a positive externality for other exporters. Exporters may band together to offer equal transfers to the importer to incentivize negotiation. Ludema and Mayda find that 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 who are unable to pay the transfer will free-ride on the eventual negotiation. In another environmental setting, Ricke, Moreno-Cruz, and Caldeira [55] also found exclusion to be optimal for coalitions deploying climate geoengineering. In their "global thermostat setting game," regional preferences lead to an incentive for more homogeneous groups to band together to enact an optimal action for the region. I confirm both of these "exclusion results" in a broader setting of negative externality reduction by coalitions with restricted actions.

1.3 Motivational Examples

In this section, I motivate the research question anecdotally and numerically. First, I portray a few real-world situations and discuss their applicability to this model. I describe this paper's notion of what a "large" actor is and what a "small" actor is. Second, I examine a simple, three-person game which previews the general result of the paper, developed more fully in the section following.

1.3.1 Anecdotal Motivation

Here, I present a few examples of negative externalities in an international context. I break down the actors of each example into "large" players and "small" players. In the context of this paper, a large player is one who takes a large action. By taking a large action, this player is the source of a large portion of the total externality. A small player, on the other hand, is one who takes a small action, possibly zero, but is still affected by the externality. This is a bit of a simplified notion of heterogeneity, which lends itself to the introductory examples I present. I expand upon this idea in favor of more nuanced heterogeneity in Section 1.5. I review the following real-world situations and label which players are large and which are small. For each, I discuss my model's applicability to the situation.

1. *Carbon Dioxide Emissions*: Consider the emission of carbon dioxide (CO_2) into the atmosphere.[4] The industry and energy provisions which release CO_2 are the action within the model, and the emissions are the negative externality.

   According to data from the European Commission Joint Research Centre, the Netherlands Environmental Assessment Agency, and the World Bank [3, 51, 61], from 2008 to 2010 China and the United States were the largest kiloton (kt) producers of CO_2 emissions, while Kiribati, Lesotho, Tuvalu, Nauru, and the U.S. Virgin Islands were the smallest producers.[5] Therefore, in terms of kiloton CO_2 emissions, China and the U.S.

[4] From the World Bank [61]: "Carbon dioxide emissions are those stemming from the burning of fossil fuels and the manufacture of cement. They include carbon dioxide produced during consumption of solid, liquid, and gas fuels and gas flaring."

[5] Both datasets indicate China and the U.S. as the largest producers but differ on the matter of the smallest. The World Bank's Development Indicators have missing values for a number of small countries, including Tuvalu and the U.S. Virgin Islands, while some

are large players, while Kiribati, Lesotho, Tuvalu, Nauru, and the U.S. Virgin Islands are small players.

In this case, China and the U.S. have large industrial sectors which manufacture goods necessary to the economy, as well as large energy demands which consume tremendous amounts of fossil fuels. The small players listed have little industry compared to the large players and in some cases no industry at all. However, since emissions affect the global stock of CO_2, the externality is felt by all players. In fact, the small island nations may be in greater danger of calamities like flooding, thereby experiencing higher expected damages from the externality [2, 40]. This example fits the model well: there is a clear distinction between large and small players, and all players benefit from reduction of the negative externality. This situation will particularly fit the type of egalitarian reduction discussed in Section 1.4.1, which prescribes the same decrease for each player.

A modification to this example is to measure the negative externality in terms of metric tons per capita, instead of kilotons of CO_2. Under this new definition of the action and according to the same data in the same time period, Qatar, Trinidad and Tobago, and the Netherlands are now the large players, since they are the top metric ton per capita producers

countries are missing entirely, like Nauru. The EDGAR database lists estimates of 1 kiloton of CO_2 for Tuvalu, Nauru, and the U.S. Virgin Islands, while the World bank estimates 18.3 kt for Lesotho and 62.3 kt for Kiribati.

of CO_2. The small players are then Burundi, Lesotho, Afghanistan, and Chad, since they have the lowest emissions of CO_2 in metric ton per capita. By this metric, China drops down to about 70th place because of its large population, despite its rank in kiloton production.

This method of measurement, however, creates a situation more difficult to describe. In a sense, the externality would result from industry and energy provision per capita, though these are odd metrics on a global scale, especially since energy use is disproportionate even within a country. Therefore, it is better to imagine the first metric (kt of CO_2) in the context of this model. If a weighted measure of CO_2 emissions is desired, then perhaps it would be more intuitive to consider emissions per industrial worker or divided by industrialized area or stock of fossil fuels, instead of emissions per capita.

2. *Albacore Tuna*: The albacore tuna (*Thunnus alalunga*) is a highly migratory species found in most of the world's oceans. The species has value as the preferred canned "white meat" tuna [7]. According to data from the United Nations' Food and Agriculture Organization [6], the highest albacore producers in 2012 were Japan, Taiwan, and China.[6] The smallest non-zero producers in 2012 were Saint Helena, Niue, Bermuda and Morocco.

   With some exception, the large albacore producers generally have large

[6] All fishing regions for a country aggregated.

fleets and advanced technology, long coasts and access to multiple fishing regions, or cooperation with other large producers. These characteristics allow for fishing multiple species, so the albacore catch is just a fraction of the total fishing business. Furthermore, the fishing industry is just a fraction of the overall economy, since these characteristics overlap with those of a rich country. The small producers generally have smaller fleets, less advanced technology for catching and processing, and small coasts. For many of the small producers, food production – and particularly for island nations, fishing – makes up a large part of the economy, and the albacore tuna is a large percentage of the industry.

> "The four species of tuna that underpin oceanic fisheries in the tropical Pacific (skipjack, yellowfin, bigeye and albacore tuna) deliver great economic and social benefits to Pacific Island countries and territories (PICTs). Domestic tuna fleets and local fish processing operations contribute 3-20% to gross domestic product in four PICTs and licence fees from foreign fleets provide an average of 3-40% of government revenue for seven PICTs. More than 12,000 people are employed in tuna processing facilities and on tuna fishing vessels. Fish is a cornerstone of food security for many PICTs and provides 50-90% of dietary animal protein in rural areas." [11]

Any voluntary reduction by the small producers greatly impacts the economy, since the fish has high marginal benefit. Any involuntary reduction, in the form of a lower catch caused by overfishing or climate change, could be ruinous. All the countries benefit from reducing overfishing caring for the health and size of the stock of albacore tuna, but perhaps the small producers perhaps benefit even more. Therefore, the global production of albacore fits the model described in this paper, because despite the heterogeneity of players, the effects of overfishing are felt by all who partake in catching this species.[7] In particular, this situation fits the type of egalitarian reduction discussed in Section 1.4.3, which respects the quickly diminishing marginal benefit of reduction by the small producers.

3. *Downstream and Downwind Pollution*: Consider two types of transboundary pollution, downstream and downwind. In each of these circumstances, there is a discrepancy in the impact of the externality on neighboring countries. Both water and air pollution can occur as byproducts of electricity production from coal, though other industrial actions can create the situation as well. An example of this case would be two countries that share access to the same river, where one country is upstream, while the other is downstream. The upstream country has waste that is released into the river to a certain extent, while the downstream country experiences the effects of this waste. In an empirical estimation testing for the presence of free-riding on water quality of international rivers, Sigman [57] finds significance for pollution by coun-

[7] There are countries who produce nil.

tries upstream of borders outside the European Union. Similarly, for an airborne pollutant, one country which is downwind of others experiences more of the externality.[8] Such conditions describe the "Black Triangle" in the Izera Mountains of the Western Sudetes [16]. In the 1980s, the forests of the Izera Mountains were decimated through the damaging combination of logging and pollutants contained in wind sediment and precipitation.

In this case, the emphasis is not on the identities of the large and small producers. Though it is likely that the upstream/upwind country is the larger producer, any heterogeneity of production could be due to different resource distributions, industrialization levels, or energy demands. What matters more in this situation is that only one agent bears the brunt of the externality, since the pollutant is quickly washed or blown away from the other. The downstream/downwind country has more to gain from reduction of the externality, and may need to compensate its neighbor.

The model in this paper allows for heterogeneity in the benefits of the action with the externality and the costs of the externality. However, this example does not fit the model as well as the previous two. The main difference is not that the externality is experienced more keenly by one party than by the others. The model allows for this, as does the first example. The real issue is that the downstream/downwind

[8] In fact, this latter situation can cause the former situation, as clouds absorb the airborne pollutant and deposit it through precipitation.

country has less of an effect on itself than the upstream country has on it. By cutting its own production, the downstream country may not meaningfully decrease the externality, nor does it affect the stock of the pollutant in the upstream country. The situation violates some of the model assumptions listed in Section 1.4.[9]

As demonstrated in these examples, the type of public-good problem described in this paper is one where all agents are affected by a negative externality, the reduction of any agent constitutes a public good for the others, and certain agents produce less of the negative externality than others. In this context, the descriptors "large" and "small" pertain to the size of an agent's action.

1.3.2 Numerical 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 no-coalition 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.

[9]Most notably, submodularity, since the upstream country is not hurt by the actions of the downstream country.

A treaty's purpose is to bring the 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 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 a specific parameterized example in a simple world of three agents, $I = \{1, 2, 3\}$, to better understand the selection of J. For this toy model, the chosen utility function is:

$$u_i(a) = \theta_i a_i - a_i \sum_{j=1}^{3} w_j a_j, \tag{1.1}$$

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

I solve first for the no-coalition Nash equilibrium, compare this to a social planner's prescriptions, 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.

In the absence of a treaty, each agent solves the following problem:

$$\begin{aligned} &\max_{a_i} \theta_i a_i - a_i \sum_{j=1}^{3} w_j a_j, \\ &\text{s.t. } a_i \in [0,1]. \end{aligned} \tag{1.2}$$

The best response function for agent i is:

$$BR_i(\theta_i, a_j, a_k) \equiv \begin{cases} 1 & \text{if } \theta_i \geq 2w_i + w_j a_j + w_k a_k \\ \frac{\theta_i - w_j a_j - w_k a_k}{2w_i} & \text{if } w_j a_j + w_k a_k < \theta_i < 2w_i + w_j a_j + w_k a_k \\ 0 & \text{if } \theta_i \leq w_j a_j + w_k a_k. \end{cases}$$

The Nash equilibrium, $a^*(\theta)$, consists of the simultaneous best responses, i.e. for each i, $a_i^*(\theta) = BR_i(\theta_i, a_j^*(\theta), a_k^*(\theta)$. Counting corner solutions, there are nine types of equilibria.[10] The interior Nash equilibrium has each agent i playing according to the function:

$$a_i^*(\theta) = -\frac{w_i \left(-4\theta_i + \theta_i w_j^2 w_k^2 + 2\theta_j w_j^2 - \theta_j w_j^2 w_k^2 + 2\theta_k w_k^2 - \theta_k w_j^2 w_k^2\right)}{2\left(4 - w_i^2 w_j^2 - w_i^2 w_k^2 - w_j^2 w_k^2 + w_i^2 w_j^2 w_k^2\right)}.$$

[10] All play zero; two play zero, one plays interior; two play zero, one plays one; one plays zero, two play interior; one plays zero, one plays interior, one plays one; one plays zero, two play one; all play interior; two play interior, one plays one; all play one.

The other equilibria involving corner solutions can be calculated. For any θ, there is only one possible equilibrium. In general, for agents with relatively large values of θ, the solution is greater than zero; I call these agents "positive producers." Meanwhile, agents with rather small values of θ take action of zero; I call these agents "non-producers." These labels are related to the concept of large and small actors, as discussed earlier.

Unless all agents are non-producers, the no-coalition Nash equilibrium is not optimal because of the negative externality. The social planner's problem is:

$$\begin{aligned} \max_{a_1,a_2,a_3} \ & \sum_{i=1}^{3} \left\{ \theta_i a_i - a_i \sum_{j=1}^{3} w_j a_j \right\}, \\ \text{s.t. } & a_i \in [0,1] \quad \forall i \in I. \end{aligned} \tag{1.3}$$

The equal-weighted Pareto optimal solution for the economy results in prescribing reduced actions for each positive producer, as expected. The threshold of producing more is now higher, meaning that an agent's benefit parameter must be very large, giving a large individual benefit, in order for an agent to be allowed to inflict a high level of externality on the others.

Having examined the no-coalition problem, I now consider a lump-sum reduction treaty for the motivating example, wherein each participant of the treaty reduces from his no-coalition 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, the minimum set of agents that must participate in the treaty, and $a^{LS}(J)$ are the agreed-upon actions.

Consider a negotiation scenario as follows: agents first determine J through some voting process, or J is somehow provided exogenously, then agents joining J choose $a^{LS}(u, J)$, while agents outside of J best respond to the actions of the coalition.[11] The outcome of the negotiation is either the singleton Nash 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 size, $\#J \in \{0, 1, 2, 3\}$.

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: with even one non-producer, the coalition of the whole can do nothing under this form of the equal treatment assumption. This pattern extends beyond the trivial case of excluding agents with corner solutions of zero and holds even for strictly interior equilibria.

In contrast, consider a different type of treaty, one of proportional reduction, where each agent participating would reduce from no-coalition 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

[11] The simultaneous timing could be altered to a Stackleberg model, where the coalition moves first in choosing their actions and the free-riders move second. The direction and exclusion results are largely generalizable, though the exact actions may differ.

still zero, albeit 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 no-coalition equilibrium. Effective treaties of this type are possible, but are not examined in this section; the toy model's focus will remain on lump-sum reduction treaties for now.

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, 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 some sort of side transfers that provide reward, 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, if one agent considers committing to reduction 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 no-coalition 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 no-coalition equilibrium. Unilateral deviation gives no outside benefit to the agent in question and allows all the other players to free-ride on the reduced action. Therefore a 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, look at the situation where agents $i = 1, 2$ are positive producers who consider the treaty:

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

while agent three best responds with $BR_3(\theta_3, a_1^* - r^*(\{1,2\}), a_2^* - r^*(\{1,2\}))$, for now suppressed as a_3^{br}. The reduction $r^*(\{1,2\})$ solves the following coalition problem:

$$\max_{r\in\mathbb{R}} \sum_{i=1}^{2} \left\{ \theta_i(a_i^* - r) - (a_i^* - r) \left[\sum_{j=1}^{2} \Big(w_j(a_j^* - r) \Big) + w_3 a_3^{br} \right] \right\}, \tag{1.4}$$
$$\text{s.t. } 0 \le r \le \min\{a_1^*, a_2^*\}.$$

With the MP commitment device, the coalition only goes into effect if the two required agents sign. Thus, agents one and two know that 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.

The comparative statics of the individual utility from this treaty can demonstrate when signing is beneficial. 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:

$$\frac{\partial u_i^{LS}}{\partial r}\bigg|_{r=0} = -\theta_i + (2w_i + w_j)a_i^* + w_j a_j^* + w_3 a_3^*. \tag{1.5}$$

This statement is positive when the benefits of reduction outweigh the foregone benefits of action, i.e. when $(2w_i + w_j)a_i^* + w_j a_j^* + w_3 a_3^* > \theta_i$. If no-coalition actions are large or θ_i is small, then this statement likely holds.

So when would the marginal utility of increasing the reduction be positive moving from zero?

1. Looking at the corner solution where all agents play an action of one, the statement clearly holds. We can rearrange it to be:

   $$(w_i + w_j + w_k) + (w_i + w_j) > \theta_i$$

   The weights add up to one, so we have:

   $$1 + (w_i + w_j) > \theta_i$$

The parameter $\theta_i \leq 1$, and the weights are strictly positive, so this statement holds. Thus when all the agents are such that they play the maximum action, reduction by two members has positive benefit. This is because the third agent, being at the maximum already, cannot free-ride upon the reduction. A treaty would be more beneficial to include him as well, but if need be, a two-person treaty is enough. This case is alluded to in Section 1.5.

2. In the case of the interior no-coalition Nash equilibrium, this condition would be:

$$(w_i(2w_i + w_j) - 2) \cdot \left(4 - w_j^2 w_k^2 - w_i^2 \left(w_j^2(1 - w_k^2) + w_k^2\right)\right) \cdot \\ \left(2w_k^2\theta_k - (4 - w_j^2 w_k^2)\theta_i - w_j^2 \left(w_k^2\theta_k - (2 - w_k^2)\theta_j\right)\right) < 0$$

   With weights of one-third for all the players, this statement becomes $17\theta_j + 17\theta_k > 323\theta_i$. We see that this is unlikely to hold for both players $i = 1, 2$ when weights on all players are equal and θ is large enough for an interior equilibrium.

   Intuitively, this makes sense, because if the equilibrium is interior, then agent three is a positive producer and has an incentive to increase his action when he is a free-rider, negating the possible benefit from reduction by the other two agents.

3. Consider the case where the third agent's no-coalition action is zero while the other two agents have interior actions $a_i^* = \frac{3}{323}(18\theta_i - \theta_j)$. Then, the

condition becomes:

$$w_j(323 + 54\theta_i - 3\theta_j) - 2(163 - 54w_i)\theta_i + 6(9 - w_i)\theta_j > 0$$

With weights of one-third for all the players, this statement becomes $19 + 9\theta_j > 48\theta_i$. It holds for many combinations of low values of θ_j and θ_i.

Furthermore, the second derivative of u_i^{LS} is negative always, so the utility of reduction is concave. Thus, if the marginal benefit of reduction is positive at zero, then the agent desires a reduction that is strictly positive.

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 have an incentive to join a lump-sum reduction treaty.

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 chosen 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 with regard to which MP constraint and vector of actions will result.

Equilibrium selection is an issue which may be resolved through game timing or bargaining protocol. A timing common to most models of the literature, such as that of Carraro et al. [20], is one where the treaty participants determine their action as a coalition in a separate stage from all agents' decision of the MP constraint. This timing reflects the idea that it would be 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 are 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 is the case in many bargaining protocols. A 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 an agreement. These ideas present areas for further research and tie-ins to other strands of bargaining literature, such as delay in negotiation and capture of bargaining position.

Using the utility function specified by Equation 1.1, 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 one's parameter θ_1 is normalized to one. The externality weights w_i are equal to one third for each agent. For each pair of parameters θ_2 and θ_3, I examine which coalitions, if any, improve the most upon the no-coalition equilibrium.

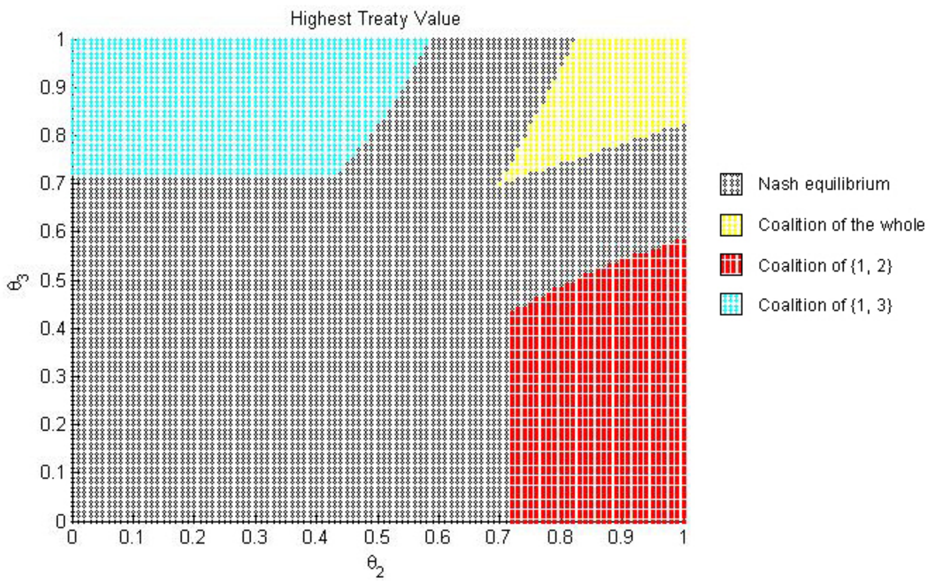

Figure 1.1: Exclusion in a linear utility function.

(a) The parameter θ_1 is normalized to one, while parameters θ_2 and θ_3 take values from zero to one. Each area depicts which coalition meets individual rationality constraint and most improves upon the no-coalition equilibrium.

For small values of the parameter, there is an area in the upper right-hand corner where the coalition of the whole is restriction Pareto optimal. However, when the parameter value decreases below some threshold, that agent drops production to zero, rendering the coalition of the whole no longer optimal under the equal treatment assumption. There is an intermediate value for the parameter where the no-coalition Nash equilibrium persists, since a coalition by the two remaining producers would be sabotaged by an increase in action from the excluded player. However, once the parameter of one agent is small enough, the remaining two form the exclusive treaty. Along the x-axis, when player two's value of θ_2 is small, there is a region where the optimal lump-sum restricted coalition is between players one and three, represented in the upper left corner. Symmetrically, along the y-axis, there is a region where the optimal coalition is between players one and two, represented in the lower right corner. The remaining region, where player one takes much larger action than player two or three, has no lump-sum treaties which improve upon the no-coalition outcome 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. In Section 1.5, I present another optimality map for a different utility function, which exhibits its own distinct pattern of exclusion.

While this toy model presented a simplified view of things, it does serve as motivation for a more generalized understanding of exclusion. In particular, it prompts the next section's result on existence of exclusion.

1.4 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.

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 continuous differentiability, 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. 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. unique Nash equilibrium, the Kuhn-Tucker conditions for equilibrium of u have a unique solution on A.

These elements describe a fairly general class of negative externality games. The main limiting assumption is that of strict submodularity. While many externality situations such as natural resource extraction fit this assumption, there are a few situations in which submodularity may be questionable, such as a firm-level analysis of emissions. A reasonable model may have that the cross-partial for a firm does not depend on the production of another firm, or that it is even enhanced by production of a complementary product. However, a strict inequality is required for the technical reason of openness and ease of proving existence; if the assumption is relaxed to no submodularity or even reversed to supermodularity, there are likely games in which the exclusion result of this holds, but this is an extension to be tackled at another point.

In this environment, I examine the possible coalitions. In a game without any cooperation, the Nash equilibria involve only the singleton coalitions.

Definition. [Nash Equilibrium.] For utility function $u \in \mathcal{U}$, the vector a^* is a **Nash equilibrium at u**, denoted $a^*(u) \in Eq(u)$, if for all players the individual vector entry $a_i^*(u)$ maximizes agent i's utility given that each other player j chose $a_j^*(u)$. Formally:

$$\forall\, i \in I,\; a_i^*(u) = \arg\max_{b_i \in A_i}\; u_i(b_i, a_{-i}^*).$$

The negative externalities condition guarantees that any no-coalition Nash equilibrium is inefficient and reductions strictly improve everyone's welfare.

Lemma 1.1. *For any utility function $u \in \mathcal{U}$, 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. Anderson and Zame [8] use a method similar in flavor in Section 4 of their paper on shyness in the proof that non-vertex pure-strategy equilibria are inefficient.

Lemma 1.1 establishes that a group of agents may form to act together. As alluded to earlier, 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. Though I discussed a few possibilities for bargaining and timing in Section 1.3.2, for the general analysis 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 no-coalition 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, a coalition's actions are certainly not even conceivable if the members do not perform as well utility-wise as in the no-coalition equilibrium; such a coalition simply would not form.

Definition. [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 no-coalition Nash equilibrium and at least one agent experiences a strict improvement. Formally,

$$(\forall j \in J)\left[u_j\left(a(u,J)\right) \geq u_j\left(a^*(u)\right)\right] \tag{1.6}$$

and

$$(\exists k \in J)\left[u_k\left(a(u,J)\right) > u_k\left(a^*(u)\right)\right]. \tag{1.7}$$

The notion of conceivability is akin to individual rationality constraints in mechanism design. Without an improvement upon her no-coalition outcome, an agent will simply not join a proposed coalition. This is also related to the unpursued issue of coalition stability: if a coalition is conceivable, then it is internally stable, since no agent within the coalition wishes to abandon it. However, I do not address external stability; I am examining coalition candidates and evaluating them in comparison to the no-coalition Nash equilibrium under different types of restricted action.

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 one-dimensional reduction from the no-coalition Nash equilibrium.

Unrestricted commitment power games for a coalition J, where J is non-empty, non-singleton, and a subset of I, are the games in which the agents in J act as a single player with the summed group utility function $\sum_{j \in J} u_j$

while every player i not in J acts as a single player with his original utility function u_i.

Definition. [Unrestricted Coalition Power.] For $J \subseteq I$, the **J-coalition game with unrestricted coalition commitment power**, denoted $\Gamma_J^{Un}(u)$, has $[\#(I \setminus J) + 1]$ agents, with the coalition labeled as agent J and having the action set $\times_{j \in J} A_j$ and 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 simultaneous-move Nash equilibrium of this game is denoted $a^{Un}(u, J) = \big(a_J^{Un}(u), a_{-J}^{Un}(u)\big)$.

As earlier discussed, Lemma 1.1 establishes that a reduction from no-coalition Nash equilibrium 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.

Lemma 1.2. *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 the actions in summed utility.*

The best possible coalition under unrestricted actions is the coalition of the whole. This is because the coalition can always request that a member plays his no-coalition action, while leaving that player out may give him leave to free-ride. If there are players with a no-coalition Nash action of zero, then coalitions excluding these players may “tie” the outcome of the coalition of the whole if the excluded players still have a best response of zero to the decreased actions of the coalition.

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. The equal treatment assumption is a broad concept, reflected in the structure of many multilateral agreements and motivated earlier in the paper. As described in the toy model, one possibility is to establish a one-dimensional decrease from no-coalition 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 [13]. This type of reduction could be particularly useful in a repeated game, where the historical equilibrium is observed and can be improved upon or reverted to.

Lemma 1.2 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, the grand 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 or proportional commitment power, the coalition of the whole may not be ideal. Sufficient heterogeneity in the costs and benefits of agents guarantees that, with restricted commitment power, coalitions strictly smaller than I are conceivable and and also improve upon a coalition of the whole, giving a sort of exclusion result. One can go even further, for some vectors of payoff functions $(u_i)_{i\in I}$, and find the set (or sets) J that are Pareto superior to the coalition of the whole amongst all subsets of I. In this section, I define two types of one-dimensional decrease from no-coalition Nash equilibrium: lump-sum commitment power and proportional commitment power.[12] I show the existence of exclusion for both.

1.4.1 Lump-sum Restricted Commitment Power

Consider the game where a coalition can commit to lump-sum reductions. By this I mean that each agent in J reduces from their no-coalition Nash action, a_i^*, by an amount r.

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

[12] Another possible restriction is an upper limit $\bar{X}$ on the stock of negative actions which is then split 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 owners' association members must restrict all noise to a lower decibel level at nighttime. This type of restriction is used in Weikard, Wangler, and Freytag [59], as well as Ludema and Mayda [48].

utility functions $u_i(a)$. The simultaneous-move Nash equilibrium of this game is denoted $a^{LS}(u, J) = \big(a_J^{LS}(u), a_{-J}^{LS}(u)\big)$.

This solution concept is subtle because the equilibrium definitions have Nash equilibria within them. This is why uniqueness of the no-coalition 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 no-coalition equilibrium as a fall-back.

The following result demonstrates that there are coalitions J, strictly smaller than the full set of agents, with lump-sum commitment power which improve upon the no-coalition 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 under the equal treatment assumption when there is enough heterogeneity.

Theorem 1.1. *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:*

$$(\forall j \in J)[u_j(a^{LS}(u, J)) > u_j(a^*(u))]. \tag{1.8}$$

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:*

$$(\forall i \in I)[u_i(a^{LS}(u, J)) > u_i(a^{LS}(u, I))]. \tag{1.9}$$

The proof relies on demonstration of a particular utility function $u \in \mathcal{U}$ for which their is exclusion, as well as openness of the conditions describing $\mathcal{U}$. The proof can be found in Appendix A. The next section elaborates on the intuition of this result using a class of models with parameterized utility functions.

1.4.2 Intuition for Exclusion under Lump-sum Restricted Commitment Power

To clarify the exclusion result given in Theorem 1.1, 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 1.4, the agents $I = \{1, 2, ..., n\}$ take 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:

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

Here, $B(a_i)$ represents the benefits of the individual action, with multiplicative coefficient θ_i. The function is increasing and weakly concave, i.e.

$B' > 0, B'' \leq 0$. 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.[13] The marginal cost function is increasing and convex, i.e. $c' > 0, c'' > 0$. Observe that since B is increasing, we have that $B'(0) > 0$, and that the cost function at zero action is zero, because $0 \cdot c(\cdot)$, so no marginal cost is incurred. These assumptions guarantee that the whole utility function is concave in all actions.

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

To further describe the heterogeneity and make use of the distinction between large and small actors, I establish two groups of players:

1. The first is group J with cardinality m.[14] The agents in group J all have $\theta_i = 1$. Hence, the utility function for players $i \in J$ is:

$$u_i(a_i, a_{-i}) = B(a_i) - a_i c\left(\sum_{j \in I} a_j\right) \tag{1.10}$$

[13] For simplicity's sake, the damages depend on the sum of the players' actions. The components of the action vector could be enter as individual arguments of the function, instead of as a sum. In a setting of incomplete information, the sum could be a discounted expectation [29].

[14] Naming this group J is an abuse of notation. It is used to suggest that this will be the group eventually forming an exclusive treaty.

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:

$$u_i(a_i, a_{-i}) = \theta B(a_i) - a_i c \left(\sum_{j \in I} a_j \right) \tag{1.11}$$

In this examination, the parameter θ increases the marginal benefit of the action as it increases from zero to one. As θ increases, the two groups grow less disparate. In equilibrium, the players in J should be taking the same action, as should all the players not in J.

Suppressing the u from the notation in the previous sections, the no-coalition Nash equilibrium of this game $a^*(\theta)$ consists of the equilibrium actions of the players not in J, denoted $a^*_{I\setminus J}(\theta)$, and the equilibrium actions of the players in J, denoted $a^*_J(\theta)$:

$$a^*_{I\setminus J}(\theta) \equiv \arg\max_{a_i \in A_i} \; \theta B(a_i) - a_i c\big(a_i + (n-m-1)\, a^*_{I\setminus J}(\theta) + m a^*_J(\theta)\big), \tag{1.12}$$

$$a^*_J(\theta) \equiv \arg\max_{a_j \in A_j} B(a_j) - a_j c\big(a_j + (n-m) a^*_{I\setminus J}(\theta) + (m-1) a^*_J(\theta)\big) \tag{1.13}$$

At the highest value of the group parameter, $\theta = 1$, both groups of players have the same maximization problem. Since the players are identical in this case, they would play the same action: $a^*_{I\setminus J}(1) = a^*_J(1)$. At lower values of the group parameter, $\theta < 1$, the two groups of agents have different maximization problems and different actions.

Lemma 1.3. *When the group parameter is strictly smaller than one, $\theta < 1$, then the equilibrium action of the players not in J, $a^*_{I\setminus J}(\theta)$, is smaller than the equilibrium action taken by players in J, $a^*_J(\theta)$.*

The proof shows that a smaller θ decreases the marginal benefit of action of the players not in J compared to that of the players in J.

According to Lemma 1.1 from Section 1.4, this no-coalition 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 no-coalition action. Examining restricted power coalitions in this class of utility functions will help illustrate the exclusion result in Theorem 1.1.

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 not in J).

First, consider the coalition of all I players. The coalition maximization problem under lump-sum commitment power is:

$$\max_{r\in\left[0,a^*_{I\setminus J}(\theta)\right]}\Bigg\{\sum_{j\in J}\Big[B\big(a^*_J(\theta)-r\big)-\big(a^*_J(\theta)-r\big)c\Big((n-m)\big(a^*_{I\setminus J}(\theta)-r\big)+m\big(a^*_J(\theta)-r\big)\Big)\Big]$$
$$+\sum_{i\in I\setminus J}\Big[\theta B\big(a^*_{I\setminus J}(\theta)-r\big)-\big(a^*_{I\setminus J}(\theta)-r\big)c\Big((n-m)\big(a^*_{I\setminus J}(\theta)-r\big)+m\big(a^*_J(\theta)-r\big)\Big)\Big]\Bigg\} \tag{1.14}$$

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 Appendix A. 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 $r_I^*(\theta) \leq a^*_{I\setminus J}(\theta)$.

For large 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 no-coalition action chosen by the players not in J, then the optimal reduction is implemented and all players participate. However, consider what happens as θ approaches zero. Then, the equilibrium action of the players not in J approaches zero as well, 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 not in J have negligible actions, the players in J continue to exert a negative externality on each other and could agree to reduce by themselves. In the limiting circumstances of θ close to zero, 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 under lump-sum commitment power is:

$$\max_{r\in\left[0,a_J^*(\theta)\right]} \sum_{j\in J} B\big(a_J^*(\theta)-r\big) \\ -\big(a_J^*(\theta)-r\big)c\left((n-m)a_{I\setminus J}^J(\theta)+m\big(a_J^*(\theta)-r\big)\right) \tag{1.15}$$

while agents not in J best respond as singletons with:

$$a_{I\setminus J}^J(\theta) \equiv \arg\max_{a_i\in A_i}(1-\theta)B(a_i) \\ -a_i\,c\left(a_i+(n-m-1)a_{I\setminus J}^J(\theta)+m\big(a_J^*(\theta)-r_J^*(\theta)\big)\right) \tag{1.16}$$

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.1), given that the players not in J were initially best responding. Even if the free-riding players increase actions from no-coalition 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 not in J have minimal response, the J coalition has $r_J^*(\theta)$ strictly greater than zero and that the exclusion result holds for small values of θ.

In the coalition of the 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 the size of that group, choosing a reduction between what would be optimal for those not in J and those in J. When the marginal benefit of action is zero for agents not in J, attempting to include them restricts the possibilities of reduction, particularly since the marginal utility of those not in J turns negative more quickly than the marginal utility of those in J.

Lemma 1.4. *There exists a threshold value $\bar{\theta} > 0$ for the group parameter such that for all values of the parameter higher than the threshold, $\theta \in (0, \bar{\theta})$, the equilibrium action of players not in J, $a^*_{I \setminus J}(\theta)$, is a binding constraint on problem (1.14).*

Lemma 1.4 demonstrates that the exclusion result from Theorem 1.1 holds for this class of parameterized 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 non-zero but small 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.

The first rationale, in particular, spurs the desire for equal treatment which still allows for small-action takers and zero-action takers to participate. This leads to the development of proportional commitment power: because a fraction of zero is still zero, even the smallest players can participate. In the next section, I examine the exclusion result under proportional restricted commitment power. General existence is once again shown, this time with the driving rationale is the disparity between the marginal benefit of reduction of heterogeneous players.

1.4.3 Proportional Restricted Commitment Power

Consider the game where a coalition can only commit to proportional reductions. A coalition J commits to a proportional reduction of s: each agent in J plays s times his no-coalition action, a proportion of what would have been played.

Definition. [Proportional Commitment Power.] For $J \subseteq I$, the **J-coalition game with lump-sum commitment power**, denoted $\Gamma_J^{LS}(u)$, has $\#(I \setminus J) + 1$ agents, with the coalition named as agent J and having the action set $\{a_J \in \times_{j \in J} A_j : (\forall j \in J)[a_j = s a_j^*(u)], s \in [0,1]\}$ and 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 simultaneous-move Nash equilibrium of this game is denoted $a^{Pr}(u, J) = \big(a_J^{Pr}(u), a_{-J}^{Pr}(u)\big)$.

The following theorem extends the exclusion result from earlier to lump-sum commitment power. Under proportional commitment power, there are coalitions J strictly smaller than I which are conceivable and Pareto-improving in the presence of heterogeneity.

Theorem 1.2. *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 $a^{Pr}(u, J)$ is conceivable, formally denoted as:*

$$(\forall j \in J)[u_j(a^{Pr}(u, J)) > u_j(a^*(u))]. \tag{1.17}$$

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

$$(\forall i \in I)[u_i(a^{Pr}(u, J)) > u_i(a^{Pr}(u, I))]. \tag{1.18}$$

The proof can be found in the Appendix. The next section gives some intuition for this result using the same class of models as in Section 1.4.2.

1.4.4 Intuition for Exclusion under Proportional Restricted Commitment Power

Recall the earlier set-up with two types of agents and separated benefit and cost functions:

1. Agents in J, or "large" agents: These agents have utility defined by Equation (1.10) and take action $a^*_J(\theta)$ in equilibrium.

2. Agents not in J, or "small" agents: These agents have utility defined by Equation (1.11), with an extra benefits parameter $\theta \in (0, 1)$, and take action $a^*_{I\setminus J}(\theta)$ in equilibrium. From Lemma 1.3, this action is smaller than the action of members in the group J.

As with the lump-sum reduction, the agents consider two possible proportional reduction treaties: one which forms a coalition of the whole, and one which contains only players in J.

For the proportional reduction coalition of all I players, the maximization problem is:

$$\begin{aligned}\max_{s\in[0,1]} \; & \sum_{i\in J} B(sa_J^*(\theta)) - sa_J^*(\theta)c\Big(\textstyle\sum_{j\in J} sa_J^*(\theta) + \sum_{k\notin J} sa_{I\setminus J}^*(\theta)\Big) \\ & + \sum_{i\notin J} \theta B(sa_{I\setminus J}^*(\theta)) - sa_{I\setminus J}^*(\theta)c\Big(\textstyle\sum_{j\in J} sa_J^*(\theta) + \sum_{k\notin J} sa_{I\setminus J}^*(\theta)\Big)\end{aligned} \tag{1.19}$$

For every θ and equilibrium $a^*(\theta)$, this will have some solution, $s_I^*(\theta)$, which lies in $[0, 1]$. This fraction solves the Kuhn-Tucker conditions in the Appendix. A choice of $s_I^*(\theta) = 1$ means that no reduction is implemented and that, in essence, the coalition agrees to play the no-coalition Nash equilibrium. On the other hand, a choice of $s_I^*(\theta) = 0$ means that the coalition implements full reduction and eliminates the negative externality and the action. Any fraction in between indicates some reduction, with lower numbers indicating more reduction than higher numbers.

As with the lump-sum reduction, the coalition of the whole works well for small values of θ. If the reduced vector, $s_I^*(\theta) \cdot a^*(\theta)$, weakly improves upon the no-coalition equilibrium, $a^*(\theta)$, for all players, then all players will participate and the coalition will implement the reduction. With this type of egalitarian treaty, there is no “physical” limit on the one-dimensional reduction choice, as with the lump-sum reduction. The lump-sum reduction was clearly limited by the smallest players’ actions; if it were larger, those players could not participate, since they could not play a negative action outside of the space $A_i = [0, 1]$. All players can physically participate, since the reduction is by a multiplied factor.

However, in revealed Section 1.4.2, there is a secondary reason why the exclusion effect holds for lump-sum restricted power. The marginal benefit of coalition reduction for a player is positive and increasing at first, then decreasing, and then negative. At some point, the reduction is too high to be optimal for that player. In fact, the reduction could go so far as to cause the players to drop out of the coalition, preferring to play the no-coalition equilibrium and then free-riding.

Furthermore, the next result shows that the grand coalition will never choose full reduction:[15]

Lemma 1.5. *For any θ, the proportional reduction taken by the grand coalition is never full-reduction, i.e.* $s_I^*(\theta) > 0$.

At first, this may seem a strong result from the problem. However, proportional reductions approaching full reduction may still occur, so the result does not limit the application in any realistic sense. In many agreements some small amount of the action is still permitted. For instance, nuclear weapons treaties allowed much of the armory already in existence to remain so, which is not a reduction to zero. For other agreements, banned behaviors may have some substitute, if imperfect: whaling provides meat and oil, which can be obtained from other animals and energy sources. The question addressed in this paper is not whether coalitions can or will reduce to zero; the question is

[15] So long as there are players with positive action – this is discussed in the Appendix.

whether a smaller coalition can improve upon the outcome of the coalition of the whole.

Consider what happens as θ approaches zero. The action of the smaller players approaches zero. At these small actions, players not in J *could* still participate in a coalition of the whole: whatever positive percentage is chosen, they can play $s \cdot a^*_{I\setminus J}$ without issue. The problem lies with whether they *would* want to reduce any further. If θ is very small, then the marginal benefit at that action increases quickly upon reduction due to concavity. Thus, agents playing a small positive action would appreciate reduction by other players, but would cling to their own last production. At some level of θ, the coalition of the whole cannot reduce at all.

Lemma 1.6. *There exists a threshold value $\bar{\theta} > 0$ for the group parameter such that for all values of the parameter lower than the threshold, $\theta \in (0, \bar{\theta})$, the reduction chosen by the coalition of the whole, $s^*_I(\theta)$, is equal to one.*

The players not in J have such small actions that any further reduction causes them great pain. Thus, consider a proportional reduction coalition just for the players in J. The maximization problem of the J-coalition is:

$$\max_{\hat{s}\in[0,1]} \; m \left[B(\hat{s}a^*_J(\theta)) - \hat{s}a^*_J(\theta)c\Big(m\hat{s}a^*_J(\theta) + (n-m)a^J_{I\setminus J}(\theta)\Big)\right]$$

The best response of those not in J is defined as:

$$a^J_{I\setminus J}(\theta) \equiv \arg\max_{a_i\in A_i} \theta B(a_i) - a_i c\Big(a_i + \textstyle\sum_{j\in J} \hat{s}a^*_J(\theta) + \sum_{k\notin J\cup\{i\}} a^J_{I\setminus J}(\theta)\Big)$$

This will have some solution, defined $s^J(\theta)$, which solves the Kuhn-Tucker condition listed in the Appendix. At any θ, the J-coalition's marginal utility of reduction from the no-coalition equilibrium is strictly positive. What this means is that, given that the players not in J were best responding by playing the no-coalition action, the benefit to decreasing from $s = 1$ is positive, i.e. $-\left(\frac{\partial u_J}{\partial s}|_{s=1}\right) > 0$ (shown in Appendix A.3.2). Therefore, if the coalition of the whole is unable to reduce because the agents not in J have very small actions, agents in J will prefer the J-coalition to form, provided that the remaining players have limited free-riding increases.

Thus, Lemma 1.6 establishes the rationale behind Theorem 1.2. Since the members of J would like to establish their own coalition, they will do so – and since the other players' actions are minuscule, their free-riding will also be negligible. Thus, the exclusion result holds for another type of one-dimensional decrease from no-coalition equilibrium.

1.5 Robustness to Policy

I have already described some reasons why restricted action sets caused by the equal treatment assumption are interesting and relevant. Here, I discuss the two chosen types in particular, placing them in context of each other and applications. I also address two alternative treaty structures related to the egalitarian restriction, the possibilities of multi-level coalitions and central commissions. Before those topics, however, I present a second optimal treaty map. While it indicates a similar pattern to the example in Section 1.3.2, it has a completely new optimal area to describe and explain.

1.5.1 Exclusion of Large Agents

So far, I have placed exclusion in the context of large actors forming coalitions without small actors. In the toy model, the only coalitions that formed included agent one, the player who always had $\theta_1 = 1$. In the proof, I present a parameterized case that splits two groups. However, Theorems 1.1 and 1.2 are more general than this. The result is general existence of exclusion, not existence hinged on this particular bifurcated heterogeneity.

Using a utility function from the set $\mathcal{U}$, I perform the same three-agent analysis as in Section 1.3.2. The utility function for agents $I = \{1, 2, 3\}$, presented below, is the one used in the Appendix in the proofs of the Theorems:

$$u_i(a) = \theta_i (10 + a_i) - a_i^2 \left(3 + \sum_{j=1}^{n} a_j \right).$$

Every set of parameters for this function has a unique interior no-coalition Nash equilibrium. For each set of parameters, I calculate the lump-sum reduction for each possible coalition and then examine which has the highest total welfare. The following image presents which coalition is optimal under the equal treatment of lump-sum reduction in this game.

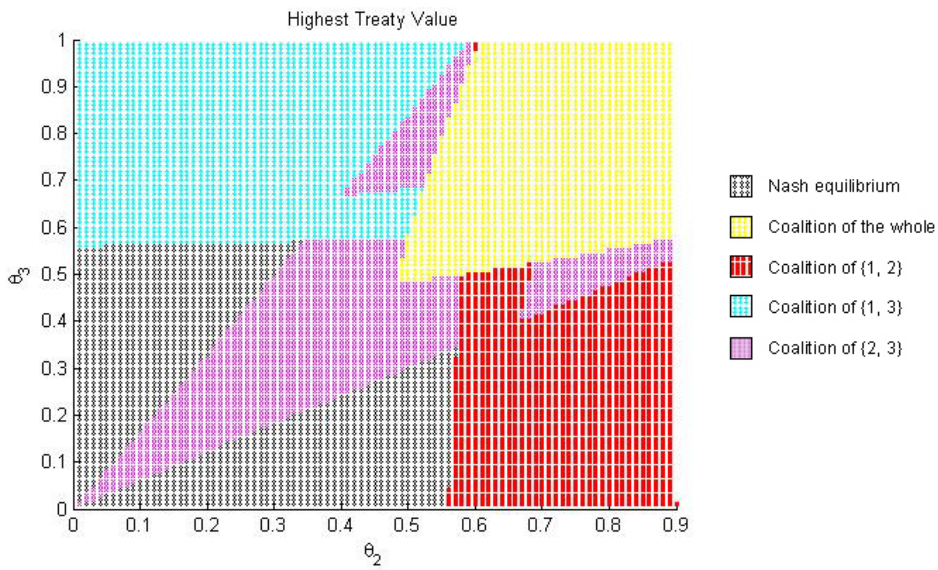

Figure 1.2: Exclusion in a parabolic utility function.

(a) As in Figure 1, the parameter θ_1 is normalized to one, while parameters θ_2 and θ_3 take values from zero to one.

The figure has a similar pattern to that of Figure 1, in that the corners match up between the two images. However, there is a whole new area where the coalition of agents two and three is optimal. This is despite the fact that agent one is normalized to be the "largest" agent (the agent with the highest marginal benefit and action). However, with this utility function, there is a low elasticity of response, and even though agent one is excluded, his action does not shift very much when he is permitted to free-ride. This creates an area where the smaller agents have the optimal coalition.

This example demonstrates the need for further examination of exclusion in treaties. While the general result has been established, there is room for more characterization regarding how and when exclusion appears. Apart from the economic models, the way exclusion applies in the context of policy must also understood, which leads to the discussion in the next section.

1.5.2 Egalitarian Restrictions as Policy

An egalitarian restriction upon actions might be an attractive idea to policymakers. First, the framing effect of "Everyone is contributing the same amount" could have a psychological impact and ease negative responses. Similarly, equal treatment serves as an easy anchoring point in negations. The restriction could result from a transaction cost to the dimension of the bargaining space. If there is some sort of cost or difficulty to bargaining over the J-dimensional action vector, then it might be easier to bargain over a one-dimensional number. Finally, the restriction guarantees a high rate of reward from the minimum participation constraint. Under the lump-sum restriction, a participant's contribution is multiplied many times over, while the proportional restriction ensures that the participants' total stock is decreased by some percentage.

The lump-sum reduction may seem too simplistic when first described, particularly for an international context of pollution or fishing reduction.[16] However, it is not so outlandish for similar entities to immediately agree to contribute the same amount, perhaps by splitting some desired total, instead

[16] A Total Allowable Catch (TAC) seems to be a reverse of the lump-sum reduction – it's a lump-sum limit. The reduction would be the difference between the original fishing level and the TAC.

of spending time haggling over exact contributions. Therein lies the exclusion: since everyone needs to contribute the same amount, the whole operation is limited by the smallest ability.

The proportional reduction, in contrast, has more real-world traction. For instance, it is easy to imagine its presence in pollution reduction treaties, where each participant has to cut emissions to, say, 80 percent of previous levels. Though the actual number which is negotiated does not depend on heterogeneity, the final contribution does. Participants who have larger actions then have a larger prescribed reduction, while participants with smaller actions still contribute some. A proportional reduction agreement has the potential to include more players, but even here, the contributions of the small actors are negligible. These players have small actions because they have a lower marginal utility of action; thus, when permitted to free-ride, they will hardly increase actions in a noticeable way.

1.5.3 Alternative Treaty Structures

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 actors may be given one lump-sum reduction, while small actors are given another. In the case of the parameterized class examined in the toy model, where much of the heterogeneity stems from a group parameter, a multi-level coalition could be a way to include the players not in J in a coalition of the whole even when θ is close to zero.

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.

Another possible structure for a treaty is to establish a central body, a commission, a third party which investigates the situation and provides recommendations. For instance, one such commission is the Inter-American Tropical Tuna Commission, which includes largest and smallest tuna producers described in Section 1.3.1 on the same committee. Such a provision can allow for some type of treaty, even if the original negotiation game is played under incomplete information and ensure longevity of an agreement. Each coalition participant then agrees to an equal treatment of sorts – supporting the establishment of such a commission – and then agrees to further action recommendations in the future. The agreement can then allow for more inclusion, perhaps overriding the benefit of treaty exclusion presented in this paper. This treaty structure merits further investigation, particularly under heterogeneity.

1.6 Conclusions

Under the restriction of egalitarian action sets, 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 unambiguously benefit the treaty. With more heterogeneous agents in the world, agents on the interior 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 a MP constraint strictly smaller than n, rendering the most disparate agents non-pivotal. In an application to environmental treaties, large polluters could use MP constraints to exclude smaller polluters and form a better performing MEA. This result is also evidenced by the use of smaller negotiation spaces: a negotiation at an exclusive summit, as opposed to the United Nations' headquarters, immediately excludes uninvited countries.

The equal treatment assumption leaves open 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 effective egalitarian treaty. Furthermore, there is some residual inefficiency to this egalitarian approach, as can be seen throughout the examples of this paper. In the numerical example of Section 1.3.2, the agents are unable to form a three-person coalition when the third agent produces zero, but the remaining agents cannot form a two-person coalition either and must play no-coalition Nash equilibrium in order to keep the third agent from free-riding on the public good. A coalition with unrestricted commitment power could prescribe an action of zero for the third agent, and an effective treaty would be adopted. This example illustrates the very tension between Lemma 1.2 and Theorems 1.1 and 1.2 – the lemma specifies that a coalition strictly smaller than the coalition of the whole cannot improve upon everyone, while the theorems give exactly the opposite. The restricted actions and differing marginal benefits of reduction drive the exclusion result. For the extreme two type set-up, exclusion gives a coalition arbitrarily close to efficiency, though this may not extend to less drastic heterogeneity.

In this model, agents are indifferent between ineffective treaties and no treaty at all, allowing symbolic agreements as equilibrium behavior. Agents could choose a lump-sum reduction of zero or a proportional reduction of one and enact an "agreement" to play Nash equilibrium. 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.

The exclusion result is hopeful, not stark; after all, the large and small producers alike prefer effective, exclusive treaties to symbolic, inclusive ones. 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. Moreover, 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.

Chapter 2

When is Bad "Bad Enough"? A Framework for Analyzing Benefits of Coordination under Externalities

2.1 Introduction

As with other resources studied in economics, international cooperation may be limited. There is only so much national effort to expend in the pursuit of negotiation with other countries, whether measured in diplomats' man-hours, dollars spent on transfers, or implementation costs. Scarcity limits the situations for which coordination can occur.

In an analogue to this idea, a person's time to haggle prices is limited. As a society, well-established super-markets do not allow for negotiation, but in the markets for cars and houses,[1] negotiation is generally expected. Produce is generally a low-stakes purchase – at least in the short term, it hardly matters if a red pepper costs \$0.99 or \$1.29 – while cars and houses are long-term financed purchases for which each dollar of price may yield much more interest over time.

[1] And even then, some people do not wish to haggle over those either, as seen in the rising popularity of car dealerships with no-haggle policies, like CarMax.

This, however, is not a paper on attrition in haggling or optimal bargaining over multiple sets. Rather, in considering international cooperation as a limited resource, a framework is desired for determining which situations are best, most important, or most beneficial to entertain for negotiation. There is an abundance of global externality situations that could benefit from an international treaty, but if there is a cost to cooperation – perhaps the opportunity cost of other things that cannot be negotiated over – then it is vital to know when a situation is more valuable cooperatively. The greatest international benefit would come from fixing the worst externality problems.

What does it mean to say an externality is "worse" in one situation than in another? This may be an easy question if social cost is measurable: producing chemical A has a social marginal cost of 5, while chemical B has one of 10. On the other hand, the question may be more difficult to answer if the externality is affected by some other parameter or construction of the situation. What happens when part of the story cannot be classified so easily? When the reason behind the severity of the externality is characterized best as something not affecting utility directly, but only through the actions chosen? For instance, if the only externality lies in a decreased chance of returning to a good state, how can that be quantified?

This paper presents a framework for the analysis of these questions, as well as sufficient conditions for a situation in which externalities are worse, based on an increasing disparity in actions between coordination and lack thereof. As a preview, this paper finds that an acceleration in the externality caused by an opponent's action, or having a large positive opposite derivative, will increase the likelihood of coordination.

Section 2.2 describes the literature surrounding this problem. In Section 2.3, I examine how to simplify a severity parameter in a one-stage game. I first discuss a symmetric game in Section 2.3.1 and then an non-symmetric game in Section 2.3.2. Chapter 3 extends the result from the previous sections to a dynamic setting by applying the framework to fishery growth models and presents conclusions.

2.2 Literature

The environmental literature has long examined international coordination on reducing negative externalities. Many externality problems, including most environmental situations, have a dynamic component to the story, beyond a one-stage fixed or marginal cost, or a tragedy of the commons. Analysis often predicates upon modeling particular situations wit some detail, such as location and travel hindering resource extraction [32], the development of technology with complementarities driving economics growth [19], or positive spillover effects increasing efficiency after increases in environmental regulation [35].

There is great interest in understanding environmental policy. Bernard and Vielle [12] use a general equilibrium model to examine dynamic climate change policy. They use the General Equilibrium Model of International National Interaction for Economy/Energy/Environment (GEMINI-E3). They attempt to calculate the closeness of a carbon tax to marginal abatement cost. They describe how with a representative consumer, this is relatively easy, but with multiple types of consumers, some suffer a welfare loss as opposed to gain. Overall, they try to examine the claim that cost of implementation of carbon tax or such may be negative and therefore there would be a "double-dividend." They find that long-term estimation works well, but that short-term analysis may be off.

Beyond these examples, there are general rules to be characterized in that the mismanagement of a shared resource with dynamic growth causes an externality to others partaking in the stock. Thus, understanding management of resource stocks aids not only the choice of an optimal time path, but also international relations. Clemhout and Wan [24] analyze two-player equilibria of harvesting continuous-time dynamic resource stocks. Under the assumptions of an exact dynamic guiding function, particular formation of performance indices, strategy spaces that are nonnegative, bounded, and locally Lipschitz, and assumed structure of coefficients, the authors prove existence of equilibrium strategy harvest plans. They then examine comparative statics for two examples: single species with stochastic evolution and two species with deterministic evolution. In particular, they examine the effects of crowding, impatience, and predator-prey relationships, looking for cases in which more is harvested. They find extra harvesting when crowding is more inhibitive to growth, when players are more impatient, and when players prefer a higher ratio of harvest to resource. In these three situations, communication and management are potentially more important for the agents.

A number of authors examine the optimal development of fish stocks, using both traditional methods of stock assessment [52, 56] and more innovative approaches [14]. A few develop structural models of dynamics caused by fish characteristics, ecological variables, and market behavior made worse by repetition, like Fischer and Mirman [32] who examine the sources of externalities in fisheries. They model and discuss the tragedy of the commons, a biological externality, a dynamic externality, and a market externality. To do so, they establish a game with two agents and two species of fish with possibly interdependent biological growth rates. Both players consume both types of fish, but only catch one species. The authors calculate the closed non-cooperative and cooperative equilibria, look at the comparative statics for both, and then compare the two to capture the market externality. In the non-cooperative equilibrium, they find that an increase in the reproductive capacity of a country's own fish species leads to a lower catch ratio due to investment value. Meanwhile, the effect of an increase in reproductive capacity of the other country's species depends on the species cross-effects. If the species have a symbiotic or negative interaction (i.e. both prey upon each other), then this causes a lower catch ratio as well, but if the species have a predator-prey relationship, then the increase in reproductive capacity of the other country's fish results in a higher catch ratio. Furthermore, making the species more symbiotic by increasing a positive cross-effect leads to a lower catch ratio; decreasing a negative cross-effect leads to a higher catch ratio;

and if there is a predator-prey relationship, it results in a lower catch ratio for the predator, a higher ratio for the prey. Clearly, understanding the relationship of the two species of fish gives insight into how the negative externality works and how to coordinate to reduce it. Compared to the market equilibrium, the cooperative catch ratio is lower under positive interactions, higher under negative interactions, and higher for a predator while lower for prey. Thus, depending on the biological parameters of the story, the benefit to coordination can change. They also find that the noncooperative equilibrium is efficient under no common access when there are no biological externalities or when the preferences are the same across countries, caused by the countries managing their own stocks separately and then selling at the fair market price.

Following along the lines of fish stocks under markets, Datta and Mirman [27] examine the interdependence between market clearing prices and harvesting decisions. They look at the entire dynamic equilibrium trajectory of each market approach and compare sources of the externality from a commons effect versus a market power effect. They find that inefficiency from overexhaustion caused by common access dominates the general restriction of market power. Their model echoes that of Fischer and Mirman in the set-up of two types of players and two types of fish. In a Cournot-Nash equilibrium, the planner in each country decides the amount of fish to be caught, taking the other countries' catch functions as given. The trading decisions and market clearing prices depend on the catches of all the countries. If there is only one country of each type, the result on efficiency of markets from Fischer and Mirman is confirmed. With more than one country for at least one type, there is dynamic overexhaustion. The authors then compare Cournot-Nash equilibrium to cooperation and to a price-taking equilibrium. They find that price-taking still causes overharvest, but performs better than Cournot-Nash under certain conditions.

In addition to the question of resource-stock management, Farzin [31] examines environmental stock externalities with threshold effects. He defines two main categories of externalities: flow externalities and two types of stock externalities. Flow externalities refer to direct damage from resource use. Resource stock externalities push up extraction cost at future dates, while environmental stock externalities cause damage from adding to accumulated stock and passing a certain threshold. First, Farzin establishes that existence of a steady-state policy requires separability of arguments in the environmental damage function and the resource extraction cost function and linearity in one of the arguments. Then using model simulation with calibration of parameters, he calculates an optimal carbon dioxide control strategy as a benchmark, finding that optimal policy postpones climate change for 122 years. The optimal policy involves no delay in implementation, as "even if for an initial period there is going to be no pollution stock damage, the optimal policy still requires that abatement begins immediately and at increasing rates"; for instance, a delay in implementation of 10 years results in over 1.2 percent loss of welfare. Since the main application is fossil fuels, when Farzin examines alternative policies, he finds that a carbon tax on its own is more effective than a sole fuel tax.

With many of these situations, parameters that change the story cause a drastic effect on the externality. Apart from Farzin's discussion of the threshold effect, a common concern in the literature is catastrophe. A catastrophe is often modeled as an uncertain, and unlikely, event which has an extremely negative impact on utility. Weitzman [60] explores the use of fat-tailed distributions to analyze a climate-change model. He finds that, in contrast to a thin-tailed model, the fat tails "can dominate the social-discounting aspect, the pure-risk aspect, and the consumption-smoothing aspect." The fat tail distribution, however, renders problems more difficult to solve.

Motoh [49] analyzes a standard resource management problem with uncertainty with the added interest of catastrophic risk. His main application is forest management with some probability of a forest fire, though the model could extend to health shocks in animal populations and pollution spillovers. Catastrophe is modeled as a Poisson value shock. Motoh finds that an increase in the risk of catastrophe, through an increase in the intensity of the Poisson process, increases a manager's optimal use rate. Because the manager is risk-averse, he prefers to use the resource more quickly than to wait for disaster to strike. Motoh's conclusions give a bit of a grim view of unpreventable catastrophe, in that the strengthening of one issue – the possibility of catastrophe – increases the problem of overharvesting in a rational way.

With regard to management of catastrophes across disparate areas, Charpentier and Le Maux [21] examine a model of the insurance market which allows for more generality applicable to a world with natural catastrophes. Breaking traditional insurance market rules, they allow for an insurer to become insolvent, as well as for correlated claims possibilities, more accurately reflecting the nature of catastrophes. They also examine whether government intervention is optimal in a single region model and in a multiple region model, each with representative consumers. In the single region, they find that the government can offer an unlimited guarantee of payment, even if the insurer falls insolvent, which increases customers' willingness to pay over the limited liability scheme offered by an insurer. With multiple regions, the authors find again that the unlimited guarantee of payment is preferred, but that there could be an issue in getting the safer region to participate in pooling of risk. The authors note that perhaps a lower premium for the safer area would help. Allowing for sovereignty of regions parallels countries in a treaty and emphasizes the need for examining such accommodations.

Accounting for quirks in the externalities proves quite important. Lange, McDade, and Oliva [45] use a catastrophe model to model technology adoption under network externalities and find that the catastrophe structure accounts much better for the adoption of PCs and PC software from 1988 to 1994. In their model, when a certain parameter is low, then small changes in an independent variable lead to small changes in the dependent variable; when the parameter value is high, then small changes in the independent variable can lead to large jump discontinuities in the dependent variable. One could model a stock externality for an environmental problem in a similar manner and search for general lessons in accounting for such parameters.

Another way of thinking of catastrophe is that prevention severely increases the benefit of coordination. In this paper, I discuss a related notion, which is the speed of increase of marginal benefit of coordination. I discuss the importance of the second derivative of an agent's utility function with respect to the opponent's action, and how it affects room for coordination. In a way, this has the feeling of Farzin's threshold effect and the catastrophe models. However, this notion differs from the literature in that it is a curvature assumption, not a probabilistic event.

The next section presents a one-stage model with an externality. I allow for "worsening" of an externality through a parameter, θ, and then discuss how changes in this parameter affect the difference between non-coordination and coordination.

2.3 Model

The first setting to examine is the simplest, and it will give basic intuition for further sections of this paper. Here is a one-shot, two player game in which players take an action which exerts an externality. For now, this is a negative externality, though Chapter 3 gives an example with a positive externality. Future work includes full extension to positive externalities. I leave the exact story, timing, and utility outcomes vague at the moment, since the goal is to describe the most general setting first and investigate individual examples afterward. I formally define the game $\Gamma = \{I, \{A_i\}_{i \in I}, \{w_i\}_{i \in I}\}$ with the following:

1. agents $I = \{1, 2\}$;
2. actions $a_i \in A_i$;
3. utility functions $w_i(a_i, a_j; \theta) \in W$, which have the properties of
   (a) twice-differentiable continuity, $w_i \in C^2$,
   (b) concavity with respect to own action, $\frac{\partial^2 w_i}{\partial a_i^2} < 0$,
   (c) negative externality, $\frac{\partial w_i}{\partial a_j} < 0$,
   (d) submodularity, $\frac{\partial^2 w_i}{\partial a_i \partial a_j} < 0$, and
   (e) unique Nash equilibrium; and
4. a "worsening parameter" $\theta \in \Theta$.

Submodularity is for convenience of the analysis at the moment, though this can be relaxed in the future. Unique Nash equilibrium allows for ease of examination. Finally, some example of functions and worsening parameters are:

1. **Fishing Boat 1.** Consider a model of a fishing boat, where a_i is effort that yields a marginal benefit depending on total actions exerted and which has a constant marginal cost:

$$w_i = a_i(1 + \theta)v(a_i + a_j) - c \cdot a_i.$$

In this example, θ multiplies the value of the action. In particular, $v(a)$ is decreasing, while a_i and a_j are perfect substitutes, so they enter certain functions as a sum; increasing either action decreases the marginal benefit of all action. A larger θ means that action is more valuable, and so intuitively, both players would increase actions and thus further diminish $v(\cdot)$.

2. **Fishing Boat 2.** Another possibility is:

$$w_i = a_i v(a_i, (1+\theta)a_j) - c \cdot a_i.$$

This is example is similar to the one above in set up, other than the fact that θ now directly magnifies the effect of the opponent's action, decreasing $v(\cdot, \cdot)$. However, unlike the example above, there is no compensating benefit from θ, and it appears that both players will lower their actions, which then gives room for compensation.

3. **Variance Spread 1.** Now consider a dynamic utility function where θ determines the entrance and effect of shocks:

$$w_i = (1-\beta)a_i[v(a_i + a_j) - c] + \beta V((s - a_i - a_j)((1-\theta)r + \theta h_t)).$$

In this example, there is a dynamic stock which affects value next period, as well as a parameter h_t carried around which affects the variability of next period's input. As θ gets larger, there is less weight on the static growth rate r and more weight on the series of h_t.

4. **Variance Spread 2.** Another possibility is:

$$w_i = (1-\beta)a_i[v(a_i+a_j)-c]+\beta E\left[V((s-a_i-a_j)((1-\theta)r+\theta h_{t+1}))|h_t\right].$$

where $h_{t+1} \sim f(h_t)$. This example is similar to the one above, except that h_t is not a known sequence. Agents can no longer perfectly prepare for what will happen, and as θ gets larger, there is more uncertainty.

5. **Time Correlation 1.** A final example is:

$$w_i = (1-\beta)u_i(a_i, a_j, s) + \beta E\left[V(a_i, a_j, s, h_{t+1})|h_t, \theta\right].$$

where $h_{t+1} \sim f(h_t, \theta)$. In this final example, which is a further extension of the ones above, θ is not even in the utility function directly, but rather governs the distribution of some shock. If this is a correlation parameter, as in Chapter 3, then this could enhance a dynamic externality.

For now, the problem will remain general. An individual player's Nash optimization problem, which has the unique solution $a_i^N(\theta)$, can be written as follows:

$$\max_{a_i \in A_i} w_i\left(a_i, a_j^N(\theta); \theta\right) \tag{2.1}$$

Because of the negative externality, the Nash equilibrium is not optimal. A social planner putting equal weight on each player would choose $a^P(\theta) = \left(a_i^P(\theta), a_j^P(\theta)\right)$, which is the unique solution to the following problem:

$$\max_{a_i \in A_i, a_j \in A_j} w_i(a_i, a_j; \theta) + w_j(a_j, a_i; \theta) \tag{2.2}$$

For both of these problems,[2] there needs to be a baseline of what occurs when the parameter is zero. Only with an understanding of a baseline can a change be measured.

Definition. The **baseline utility function** of a game Γ is evaluated at $\theta = 0$, and is formally written as:

$$w_i(a_i, a_j; 0) = u_i(a_i, a_j) \tag{2.3}$$

The **baseline optima** are as follows:

1. The non-cooperative Nash equilibrium is denoted as $a^N(0)$, abbreviated as $\mathbf{a^N}$, with components $\left(a_i^N(0), a_j^N(0)\right)$, which may also be abbreviated to $\left(\mathbf{a_i^N}, \mathbf{a_j^N}\right)$. This is the unique solution to the simultaneous maximization problems for all i in I:

$$\max_{a_i \in A_i} u_i\left(a_i, a_j^N\right)$$

2. The cooperative Nash equilibrium, or Social Planner's solution, is denoted as $a^P(0)$, or shortened to $\mathbf{a^P}$, with components $\left(a_i^P(0), a_j^P(0)\right)$,

[2] Observe that the first order conditions to both of these problems could be summarized as the following set of equations, where the Nash condition is at $t = 0$, while the social planner's condition is at $t = 1$:

$$\frac{\partial w_i}{\partial a_i} + t\frac{\partial w_j}{\partial a_j} = 0, \quad t\frac{\partial w_i}{\partial a_i} + \frac{\partial w_j}{\partial a_j} = 0.$$

which may also be abbreviated to $\left(\mathbf{a_i^P}, \mathbf{a_j^P}\right)$. This is the unique solution to the social planner's maximization problem:

$$\max_{a_i \in A_i, a_j \in A_j} u_i(a_i, a_j) + u_j(a_j, a_i)$$

The existence of negative externalities means that coordination, if possible, would be Pareto-improving. One of the main interests is when coordination will happen and its resulting value. It is possible that when value is higher, coordination is more likely. However, what does it mean for value to be higher? To answer this question, I examine the difference between non-coordination and coordination, taking into account gaps in utility and in action. The possible objects of interest to pursue are:

1. **Direct value to coordination:** This seems to be the clear measure of benefit: how much extra total surplus can be created by moving from non-coordination to coordination in situations with varying degrees of externality? If θ characterizes the externality, then of interest is how increases in θ, which make the externality worse, will affect the gap between coordinated and uncoordinated action:

$$\frac{d}{d\theta}\left[w^P(\theta) - \left(w_1^N(\theta) + w_2^N(\theta)\right)\right]$$

where the $w^P(\theta)$ is the total utility evaluated at $a^P(\theta)$, $w_1^N(\theta)$ is the utility to agent one evaluated at $a^N(\theta)$, while $w_2^N(\theta)$ is the utility to agent two evaluated at $a^N(\theta)$.

The difference should always be weakly positive, because of the definition of the two problems. However, the gap could stay the same as θ increases, or could even shrink. Therefore, it is of non-trivial interest to characterize when this gap is strictly increasing.

Unfortunately without careful attention to the structure of the problem, this object could capture changes purely in levels. It appears this "value to coordination" can be arbitrarily manipulated via magnitude of the gap. This leads to another object of interest.

2. **Increase in range of coordination:** In the presence of a negative externality, a social planner's actions are generally smaller than the Nash actions. Rather than pursuing changes in utility, one way to think of an externality getting worse would be if the social planner's recommended actions are decreasing as the worsening parameter increases, while agents acting on their own are inclined to do the opposite. An increasing gap between actions taken under coordination and non-coordination can be another sign that an externality is getting worse. Therefore, of interest is how the difference between the Nash equilibrium and the Pareto optimal actions change with respect to the parameter:

$$\frac{d}{d\theta}\left[a_i^N(\theta) - a_i^P(\theta)\right].$$

If the actions are moving further apart from one another, there may be more benefit to coordination. Furthermore, the scope of possible agreements is increased, and there are more reductions that can be made, so this may be another notion of when a treaty is more likely.

The above can also be written as:

$$\frac{\Delta a_i^N(\theta) - \Delta a_i^P(\theta)}{\Delta\theta}$$

Moving from an original utility function, the change in the parameter is simply the value assigned, that is

$$\Delta\theta = \theta - 0 = \theta.$$

The changes in the Nash and Pareto optimal actions can be written as:

$$\Delta a_i^N(\theta) = a_i^N(\theta) - a_i^N(0)$$

$$\Delta a_i^P(\theta) = a_i^P(\theta) - a_i^P(0)$$

Observe that were one to multiply these by θ, these would look like pieces of a first order Taylor expansion around the Nash equilibrium and the social optimum.

For certain problems, even if it is possible to determine how the parameter θ affects actions, the relative changes between coordination and non-coordination may be difficult to characterize. Furthermore, if θ is a difficult nature to derive, simplification through linearization may assist in answering the questions of interest.

In the following section, I describe a symmetric game within the basic assumptions described earlier in order to gain some intuition in a simple case.

I modify the original utility functions with three different linearized worsening parameters – an "own effect," an "opponent effect," and a "submodular effect" – in order to represent more complicated utility functions. I then derive general conclusions for such parameters in a one-stage symmetric game.

2.3.1 Symmetric Game

There are many ways to model the severity of an externality, depending on the type of influence the action has upon it. For instance, the simplest notion of worsening could consist of "pure hurt," a multiplying factor on the opponent's action which does not affect marginal utility of own action but which lowers utility unambiguously. A more complicated version of worsening could involve a story of correlation in time shocks of a resource stock, and as more information is available, the resource stock is exploited even more and the tragedy of the commons worsens.

As mentioned earlier, this section models the worsening of externalities using three paths: changing how the opponent's action affects utility, changing how the agent's own action affects utility, and changing how the cross-effect of actions affects utility. These three paths offer representation of more complicated stories on their own or through combinations.

With regard to the earlier discussion of utility gaps versus action gaps, there are a few ways to go about adding a linearized term to the utility function. One possibility is to simply add a linear term multiplied by θ. This will change the derivative with respect to that variable in a linear manner.

The "pure hurt" term would be represented as subtracting off $\theta \cdot a_j$ from the baseline utility function. This approach will give the correct intuition for the action gaps, but will necessarily affect utility as well. There is some worry that an increase in the utility gaps between non-coordination and coordination is somehow "built in" through this term.

Another approach is to model the parameter effects as if they were Taylor expansions around the Nash equilibrium or the Pareto optimal solution, so these changes in externalities can be thought of as affecting the derivatives of a symmetric utility function. These would not affect utility through the additive term unless actions changed. However, while taking a derivative at two different points is a mathematically sound idea, the economic intuition is somewhat murkier. This exercise also captures the correct directional changes in actions, but may cause concern that the utility function under coordination is different than that under non-coordination.

Because of these concerns, I use the simple linearization to study only the action gaps. I ignore the direct value to coordination, because of the limitations mentioned earlier. I do present the alternative Taylor expansion structures in the Appendix, and initial analysis for them appears to be similar.

The *opponent effect* is the most intuitive of the three linearizations. This is where the worsening of the externality rotates the first derivative with respect to the opponent's action. The linearization for the individual problem is:

$$
\begin{aligned}
w_i^N(\theta_J) &= u_i(a_i, a_j) - \theta_J a_j, \\
w_j^N(\theta_J) &= u_j(a_j, a_i) - \theta_J a_i.
\end{aligned}
\tag{2.4}
$$

The own effect linearization for the Social Planner's problem is:

$$
w^P(\theta_J) = u_i(a_i, a_j) + u_j(a_j, a_i) - \theta_J \left(a_i + a_j\right). \tag{2.5}
$$

As θ increases there is more room for an increase in the opponent's action to harm the player. Thus, as θ increases, the externality is worsening, particularly compared to level of $\theta = 0$.

The *own effect* improves the value of one's own action, incentivizing agents to take larger actions. With a submodular utility function, this enhanced activity decreases the marginal benefit of the opponent, thereby increasing the negative externality. Here, the worsening of the externality is the rotation of the first derivative with respect to own action. The linearization for the individual problem is:

$$
\begin{aligned}
w_i^N(\theta_I) &= u_i(a_i, a_j) + \theta_I a_i, \\
w_j^N(\theta_I) &= u_j(a_j, a_i) + \theta_I a_j.
\end{aligned}
\tag{2.6}
$$

The own effect linearization for the Social Planner's problem is:

$$
w^P(\theta_I) = u_i(a_i, a_j) + u_j(a_j, a_i) + \theta_I \left(a_i + a_j\right). \tag{2.7}
$$

This equation is very similar to Equation (2.5); the main difference is that the sign on the worsening parameter is opposite. When linearizing the parameter, the determination of own or opponent effect in the social planner's problem is reduced to the sign on the coefficient. The increase of the parameter θ_I increases the value of acting, which may in fact override the externality at some point, when individual benefit outweighs social cost. Thus it can be expected that changes of this kind eliminate the need for coordination at high levels, though at low parameter values there might still be benefit.

The *submodular effect* changes the cross-partial of both actions, making the utility function more submodular than before and enhancing the negative externality in this manner. The linearization for the individual problem is:

$$\begin{aligned} w_i^N(\theta_{IJ}) &= u_i(a_i, a_j) - \theta_{IJ} a_i a_j, \\ w_j^N(\theta_{IJ}) &= u_j(a_j, a_i) - \theta_{IJ} a_i a_j. \end{aligned} \tag{2.8}$$

The submodular effect linearization for the Social Planner's problem is:

$$w_i^P(\theta_{IJ}) = u_i(a_i, a_j) + u_j(a_j, a_i) - 2\theta_{IJ} a_i a_j. \tag{2.9}$$

The effect of each separate modification can be found by comparing the first order conditions of the altered coordination and non-coordination problems. The following theorem gives the direction of the action gap for each effect in a symmetric game.

Theorem 2.1. *For a symmetric game Γ, an increase in the parameter multiplying the added linearizations has the following effect for each:*

1. *Increasing the opponent effect increases the distance in the actions under non-coordination and coordination, that is, for all i:*

$$\frac{d}{d\theta_J}[a_i^N(\theta_J) - a_i^P(\theta_J)] > 0;$$

2. *Increasing the own effect has ambiguous results on the distance in actions under non-coordination and coordination; and*

3. *Increasing the submodular effect also has ambiguous results on the distance in actions under non-coordination and coordination.*

The proof of Theorem 2.1 is the Appendix, though its intuition is discussed briefly here. As mentioned earlier, the opponent effect is perhaps the most intuitive, and it is easy to see why its effect is unambiguous. Adding the linearized term to the problem of non-coordination does not change the player's own incentives, so the Nash actions are unchanged. However, this term changes the incentives facing a social planner, and coordinated actions decrease. The opponent effect could model a story where there is simply a larger harm from the opponent's action, or a more complicated story where harm from the opponent's action prevails.

The own effect is ambiguous, at least in attempting to describe it for the whole range. At small increases, it can result in a positive gap, due to the submodularity in the problem. However, the enhanced benefit from an increase in one's action at some point outweighs the increased negative externality caused by the other player doing the same. A social planner would also increase actions, but more slowly because of the negative externality and submodularity. An example of this would be an improved technology that increases the marginal benefit of own action.

The submodular linearization is also ambiguous unless curvature is examined, but for another reason. While the own effect caused increases in actions under both coordination and non-coordination, the submodular effect causes decreases in both. The direction of change in the distance of action gaps depends on the comparative speeds of reduction.

For this symmetric analysis, the three effects were examined separately, in order to distinctly characterize each. In translating an externality situation into this linearized parameter, the three effects may need to be combined to correctly capture the circumstances. This idea is further explored in the next chapter of this dissertation.

2.3.2 Non-symmetric Game

The symmetric game places assumptions on the direction that the responses to θ can take. For instance, cases where the same change in θ affects the players differently are not permitted. Thus, in examining the many ways an externality could be worse, non-symmetric games are important as well. Moving to a non-symmetric game opens up more possible outcomes with regard to direction of the players' reactions, and the directions derived in Theorem 2.1 may no longer hold.

Because there are fewer restrictions, this section will ignore the submodular effect in order to keep the analysis tractable. Using both the own effect and the opponent effect allows for the agents to affect each other asymmetrically. The parameter θ_{xy} represents a deepening of player x's effect on player y. With this adjusted linearization, the agent's utility functions are now:

$$\begin{aligned} w_i^N(\theta) &= u_i(a_i, a_j) + \theta_{ii} a_i - \theta_{ji} a_j, \\ w_j^N(\theta) &= u_j(a_j, a_i) - \theta_{ij} a_i + \theta_{jj} a_j. \end{aligned} \tag{2.10}$$

The non-symmetric linearization for the social planner is:

$$w^P(\theta) = u_i(a_i, a_j) + u_j(a_j, a_i) + (\theta_{ii} - \theta_{ij}) a_i + (\theta_{jj} - \theta_{ji}) a_j. \tag{2.11}$$

As briefly alluded to earlier, the expansions of interest have a linear combination of coefficients in front of them. However, while the symmetric game assured that both coefficients collapsed into only one, here there are two distinct coefficient. Therefore, two composite coefficients can be defined as allows:

$$\begin{aligned} \gamma_i &= \theta_{ii} - \theta_{ij} \\ \gamma_j &= \theta_{jj} - \theta_{ji} \end{aligned} \tag{2.12}$$

With these coefficients, the social planner's problem can be rewritten as:

$$\max_{a_i, a_j} \; u_i(a_i, a_j) + u_j(a_j, a_i) + \gamma_i a_i + \gamma_j a_j \tag{2.13}$$

Because the two effects linearly combine, whether there is an own effect or an opponent effect for each agent is given by the signs of γ_i and γ_j. There are five main regions of interest: both γ_i and γ_j are positive, both are negative, they are of opposite signs, one is zero while the other is positive, and one is zero while the other is negative.

If both coefficients are negative, then there is an opponent effect only. From the previous section, it is safe to say that coordination will reduce actions, while the non-coordination actions are constant or increasing. If they are of opposite signs, or one is negative while the other is zero, then it is likely that the agent causing an opponent effect will have his action reduced, while the other agent's action may be increased. Also from the previous section came the result that the own effect is ambiguous and depending on curvature. This is of great interest and will now be somewhat resolved for the non-symmetric case. The following analysis will characterize sufficient conditions for diverging actions under own effect, or in areas where $\theta_{ii} > \theta_{ij}$ and $\theta_{jj} > \theta_{ji}$.

The Nash equilibrium, a^N, uniquely solves the following simultaneous best response problems:

$$\max_{a_i} u_i(a_i, a_j) + \theta_{ii}a_i - \theta_{ji}a_j$$
$$\max_{a_j} u_j(a_j, a_i) - \theta_{ij}a_i + \theta_{jj}a_j$$

The Nash first order conditions are:

$$\frac{\partial u_i(a_i, a_j)}{\partial a_i} + \theta_{ii} \equiv 0,$$
$$\frac{\partial u_j(a_j, a_i)}{\partial a_j} + \theta_{jj} \equiv 0.$$

In the symmetric case, the direction of movement of actions could be determined because of the extra assumptions symmetricity imposed. Now, however, the actions could be moving in separate directions, as the parameters θ_{ii} and θ_{jj} can also move around separately. One similarity to the symmetric case is that the opponent effect coefficients wash out, no longer appearing in the first order conditions. Any effect from the opponent will come from their own adjustment of action. The directions of changes can be analyzed with a second order expansion. This process can be found in the Appendix. The comparative statics that result can be summarized as follows:

$$U^N = \begin{bmatrix} \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2} & \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} \\ \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j} & \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2} \end{bmatrix}$$

$$D_{\theta_{ii}} a^N = \begin{bmatrix} \frac{\partial a_i^N}{\partial \theta_{ii}} \\ \frac{\partial a_j^N}{\partial \theta_{ii}} \end{bmatrix}, \quad D_{\theta_{jj}} a^N = \begin{bmatrix} \frac{\partial a_i^N}{\partial \theta_{jj}} \\ \frac{\partial a_j^N}{\partial \theta_{jj}} \end{bmatrix}, \quad Da^N = \begin{bmatrix} D_{\theta_{ii}} a^N & D_{\theta_{jj}} a^N \end{bmatrix}$$

$$U \cdot Da^N = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

All of the entries in U are negative due to concavity and submodularity. This means that the entries in $D_{\theta_{ii}} a^N$ need to be opposite signs, as do the entries in $D_{\theta_{jj}} a^N$, in order to obtain the negative identity matrix when multiplied with U^N.

Since θ_{ii} is not in player j's first order conditions, the effect of θ_{ii} on j's action can be described as follows:

$$\frac{\partial a_j^N}{\partial \theta_{ii}} = \frac{\partial a_j^N}{\partial a_i} \cdot \frac{\partial a_i^N}{\partial \theta_{ii}}$$

Because of the submodularity, $\frac{\partial a_j^N}{\partial a_i} < 0$, and because of the own effect, $\frac{\partial a_i^N}{\partial \theta_{ii}} > 0$. Hence, $\frac{\partial a_j^N}{\partial \theta_{ii}} < 0$, so the two are of opposite signs and the story can hold under proper curvature assumptions. This idea is similar to Huang and Smith's discussion of congestion versus agglomeration and determining the direction of externalities in shrimp fishing [38].

For the cooperative problem a social planner chooses unique $a^P(\theta)$ to solve:

$$\max_a \; u_i(a_i,a_j) + u_j(a_j,a_i) + \gamma_i a_i + \gamma_j a_j$$

The first order conditions are:

$$\begin{aligned} \frac{\partial u_i(a_i,a_j)}{\partial a_i} + \frac{\partial u_j(a_j,a_i)}{\partial a_i} + \gamma_i \equiv 0, \\ \frac{\partial u_i(a_i,a_j)}{\partial a_j} + \frac{\partial u_j(a_j,a_i)}{\partial a_j} + \gamma_j \equiv 0. \end{aligned}$$

Once again, the second order expansions are in the Appendix. The summary of the comparative statics is:

$$U^P = \begin{bmatrix} \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2} + \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i^2} & \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} + \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j} \\ \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j} + \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} & \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2} + \frac{\partial^2 u_i(a_i,a_j)}{\partial a_j^2} \end{bmatrix}$$

$$U^P = U^N + V, \quad V \equiv \begin{bmatrix} \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i^2} & \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j} \\ \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} & \frac{\partial^2 u_i(a_i,a_j)}{\partial a_j^2} \end{bmatrix}$$

$$D_{\gamma_i} a^P = \begin{bmatrix} \frac{\partial a_i^P}{\partial \gamma_i} \\ \frac{\partial a_j^P}{\partial \gamma_i} \end{bmatrix}, \quad D_{\gamma_j} a^P = \begin{bmatrix} \frac{\partial a_i^P}{\partial \gamma_j} \\ \frac{\partial a_j^P}{\partial \gamma_j} \end{bmatrix}, \quad Da^P = \begin{bmatrix} D_{\gamma_i} a^P & D_{\gamma_j} a^P \end{bmatrix}$$

$$U^P \cdot Da^P = (U^N + V) \cdot Da^P = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Because of submodularity and concavity, all the entries in U^N are negative, and the off-diagonal entries in V are negative as well. As of yet, however, this paper has placed no assumptions on the diagonal entries in V. The diagonals are second derivative with respect to opponent's action, an aspect which is not commonly modeled.

In many common utility functions, the opponent-directed second derivative is zero. Once the opponent's initial effect is known, externality or not, rarely is the speed of that effect explicitly described as being central to the problem. If the second derivative is zero, $\frac{\partial^2 u_i}{\partial a_j^2} = 0$, this means that the opponent's effect is constant, and that regardless of the opponent's action, their marginal externality will be the same. For a negative externality, if the second derivative is negative, $\frac{\partial^2 u_i}{\partial a_j^2} < 0$, then the opponent's effect on utility is 'accelerating' – as the opponent's action is increasing, the marginal externality is becoming more negative. If the second derivative is positive, $\frac{\partial^2 u_i}{\partial a_j^2} > 0$, this means a negative externality is 'decelerating.' As the opponent's action is increasing, the marginal negative effect on the player is becoming less negative.[3] Borrowing the term from physics, the second derivative of distance with respect to time is acceleration; this concept gives an idea to the incentives of reduction of a negative externality or promotion of a positive externality. When the benefits to coordination are not only increasing but accelerating, negotiation is warranted.

In order to determine the worsening parameter effect on the gap between actions, the D matrices need to be understood. It can be shown that

[3]For a positive externality, a negative second derivative is decelerating the effect of the externality, while a positive second derivative is accelerating.

$Da^N = -U^{-1}$ and $Da^P = -(U+V)^{-1}$. The main questions are how these behave and where $Da^N - Da^P$ has a definite positive sign.

First, some standardization is called for. The derivatives in Da^N are in fact with respect to θ_{ii} and θ_{jj}, while those in Da^P are with respect to γ_i and γ_j, which are the composite coefficients defined earlier. In order to compare the two, it needs to be shown that the hypothetical derivative of a^N with respect to γ_i and γ_j is the same as already taken for θ_{ii} and θ_{jj}.

Lemma 2.1. *The derivative of a^N with respect to θ_{ii} and θ_{jj} is equal to the derivative of a^N with respect to γ_i and γ_j, i.e.*

$$\begin{bmatrix} \frac{\partial a_i^N}{\partial \theta_{ii}} & \frac{\partial a_i^N}{\partial \theta_{jj}} \\ \frac{\partial a_j^N}{\partial \theta_{ii}} & \frac{\partial a_j^N}{\partial \theta_{jj}} \end{bmatrix} = \begin{bmatrix} \frac{\partial a_i^N}{\partial \gamma_i} & \frac{\partial a_i^N}{\partial \gamma_i} \\ \frac{\partial a_j^N}{\partial \gamma_j} & \frac{\partial a_j^N}{\partial \gamma_j} \end{bmatrix}$$

Proof. Recall the definition of the composite coefficients:

$$\gamma_i = \theta_{ii} - \theta_{ij}$$

$$\gamma_j = \theta_{jj} - \theta_{ji}$$

When taking the total derivative of a_i with respect to γ_i, the following is obtained:

$$\frac{\partial a_i^N}{\partial \gamma_i} = \frac{\partial a_i^N}{\partial \theta_{ii}} - \frac{\partial a_i^N}{\partial \theta_{ij}}$$

It has already been obtained that $\frac{\partial a_i^N}{\partial \theta_{ij}} = 0$, hence we have $\frac{\partial a_i^N}{\partial \gamma_i} = \frac{\partial a_i^N}{\partial \theta_{ii}}$. This can be repeated for γ_j, and then for a_j^N. Thus, the two matrices are equivalent. □

In such a general setting, it is difficult to say where the derivatives are positive or negative. Therefore, instead of looking for necessity, one possible approach is to look for sufficient cases of possible direction. The action gap is certainly increasing if a_i^P decreases while a_i^N grows or remains constant, or if a_i^P remains constant while a_i^N grows. More difficult situations would involve relative speeds of the two and will remain unaddressed in this paper. Thus, this means it is of interest to figure out when Da^N is positive in both entries while Da^P is negative in both entries when both parameters are changed in the same direction, if not by the same magnitude.

First, I examine the Nash actions to find when Da^N is positive. Then, I examine the social planner's actions to find when Da^P is negative. The matrix U^P is a bit more complicated than U^N, so I will use two different approaches.

To find the responses of the Nash actions when the non-symmetric own effects are both increasing, I look for sufficient conditions for the following to be positive in both entries:

$$
\begin{aligned}
U \cdot Da^N &= \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \\
Da^N &= U^{-1} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -U^{-1} \cdot I = -U^{-1} \\
&= - \begin{bmatrix} \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2} & \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} \\ \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j} & \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2} \end{bmatrix}^{-1}
\end{aligned}
$$

When Da^N is written as $\begin{bmatrix} D_{\theta_{ii}} a^N & D_{\theta_{jj}} a^N \end{bmatrix}$, if both entries are positive, then Da^N is positive as well. This method of writing Da^N will be a linearization and can be found by the following procedure:

$$\begin{bmatrix} D_{\theta_{ii}}a^N & D_{\theta_{jj}}a^N \end{bmatrix} = -\begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2} & \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} \\ \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j} & \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2} \end{bmatrix}^{-1}$$

If the linearized inverse is negative, then the whole expression will be positive. Recall that the inverse of a 2×2 matrix is:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Using the definition of an inverse and applying the linearization, observe that:

$$-\begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} > 0 \quad \text{iff} \quad \frac{1}{ad-bc} \begin{bmatrix} d-c & a-b \end{bmatrix} < 0$$

Lemma 2.2. *For Da^N to be positive and for the Nash actions to be increasing in response to an increase in θ, it is sufficient for the own second derivatives to be the same direction in comparison to the cross-partials for both agents. That is, the own second derivative can be more negative than the cross partial for both agents:*

$$\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2} < \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} \quad \text{and} \quad \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2} < \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j}$$

or, the own second derivative can be less negative than the cross partial for both agents:

$$\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2} > \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} \quad \text{and} \quad \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2} > \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j}$$

The proof of Lemma 2.2 is in the Appendix. Having found sufficient conditions for Da^N to be increasing, I now examine the movement of Da^P, the actions under coordination. As before, observe that:

$$\begin{aligned} Da^P &= -(U+V)^{-1} \\ &= -\begin{bmatrix} \frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} & \frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j} \\ \frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j} & \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2} \end{bmatrix}^{-1} \end{aligned}$$

Since $(U+V)$ is symmetric, its inverse is also symmetric. Furthermore, a symmetric matrix is diagonalizable, so the eigenvalues of the inverse matrix can be used to figure out the sign of its determinant. Since the matrix is diagonalizable, there is some Q such that:

$$(U+V)^{-1} = Q^T \Lambda Q$$

Since Q is repeated, the sign of the expression is determined by Λ, the matrix of eigenvalues. If the eigenvalues of $(U+V)$ are positive, then the eigenvalues of its inverse will be as well. When multiplied by the outside negative, Da^P will be negative, providing the decreasing effect desired.

Lemma 2.3. *For U^P to have only positive eigenvalues, it is sufficient that:*

$$\frac{\partial^2 u_i}{\partial a_j^2} > 0 \tag{2.14}$$

$$\frac{\partial^2 u_j}{\partial a_j^2} > 0 \tag{2.15}$$

$$\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) > \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2 \tag{2.16}$$

Lemma 2.3 is proven in the Appendix. Combining it with Lemma 2.2, the following theorem is obtained.

Theorem 2.2. *For a symmetric or non-symmetric game* Γ*, it is sufficient for a utility function to satisfy Lemmas 2.2 and 2.3 in order for an increase in the parameters multiplying the added linearizations to increase the distance in actions under non-coordination and coordination.*

Proof. If Γ satisfies Lemma 2.2, then Da^N is positive, so a^N is increasing in θ_{ii} and θ_{jj} and unresponsive to θ_{ij} and θ_{ji}. If Γ satisfies 2.3, then Da^P is negative, so a^P is decreasing in γ_i and γ_j, while γ_i is increasing in θ_{ii} and decreasing in θ_{ij} and γ_j is increasing in θ_{jj} and decreasing in θ_{ji}. $\square$

With Theorems 2.1 and 2.2, this chapter has established two interesting cases of sufficiency of increasing action gap: the opponent effect only in a symmetric game, and the own effect in a possibly non-symmetric game under certain curvature assumptions. The next chapter describes how the framework can be applied and attempts to do so in the context of a fishery.

Chapter 3

Applying the Framework to Fishery Models

3.1 Introduction

As the previous chapter demonstrated, an indirect externality parameter can drive whether a situation merits international cooperation. A framework was developed for approximating complicated externalities with a linearized parameter or combination of parameters. The current chapter now applies the framework to fisheries and a dynamic externality caused by growth correlation.

There is much concern regarding the state of the world's fisheries today. Current harvesting strategies have caused great depletion in ocean stocks. Unlike land animals, the actual stock of a species of fish can be extremely difficult to assess. Estimating the population of a fish stock requires understanding of three dynamic rates: recruitment, or the rate at which a juvenile fish is consider mature enough to be caught; individual growth, the rate at which members of the species grow in length; and mortality, the rate at which fish die from both fishing and natural causes. Hence, "[u]nless the rate of harvesting can be controlled somehow, the fish population may eventually be reduced (at a profit) to a low level. This in turn may affect the productivity of the resource and greatly reduce future catches" [23]. There is evidence that humans are "fishing down the food web," seen in a declining average trophic level of worldwide catches [53], which indicates unsustainable fishing strategies.

By their very nature, ocean fish are an international commodity, and so any meaningful coordination among producers must come through an international agreement. There are 112 accords, agreements, conventions, protocols, and amendments under the keyword "fisheries" in the Environmental Treaties and Resource Indicators database, part of the Socioeconomic Data and Applications Center hosted by the Center for International Earth Science Information Network at Columbia University [33]. About 10 of these have "tuna" in the title. There are number of species of tuna which are "among the most valuable commercial species," valued for their taste and sportiveness [41]. None of the 112 agreements, however, have "shrimp" in the title. Yet both species are especially valuable; "[i]n the United States, the annual landings of tuna are usually surpassed in monetary value only by the shrimp catch" [41].

Tuna and shrimp have very different reproductive patterns. Tuna are classified as highly migratory species, with some like the albacore tuna making trips from California to Japan. Spawning females release "100,000 eggs per kilogram" of their body weight, and some mega-spawners can weigh 65 kilograms. Larva and juvenile mortality rates are high, but "[t]o keep the tuna population constant, only two offspring from the millions of eggs produced by each female would have to survive to maturity" [41]. This pattern suggests high time dependence, where the amount of tuna greatly depends on the number of fish of recruitment age.

Shrimp, on the other hand, hug the coastline. “Juvenile shrimp migrate inshore in the spring, grow in estuaries during the summer and fall, and then swim back to the open ocean to spawn in the winter and spring. This behavior results in a major harvest season from early summer to early winter that concentrates in the estuaries and nearshore in the open ocean” [38]. Compared to tuna, this pattern suggests a model where recruitment is less vital year-to-year.

While externalities in fishing will arise from multiple sources, including crowding and commons problems, this chapter applies the framework developed in the previous to the dynamic externality resulting from changes to the good growth periods of fish. Section 3.2 presents some of the biological and economic literature of fish growth. Section 3.3 presents a simplified model of a dynamic fishery and describes a parameter which causes an increasing action gap, while Section 3.4 develops a more realistic fishing model and demonstrates how the framework would be applied. Section 3.5 gives conclusions for these two related chapters.

3.2 Literature

Individual fisheries and governments rely on stock assessment techniques, particularly those using biomass approximation, to estimate the number of fish in the oceans.[1] By monitoring data on catches, abundances, and biological relationships, it is possible to estimate the size of a fish stock [1], as well as important variables including recruitment growth [56], individual growth, seasonal growth [25], and additional environmental interactions [17, 36, 50].

Biologists and economists also analyze human impact of harvesting strategies and market behavior on fish populations [46]. Cabral, Geronimo, Lim, and Aliño [18] examine how a two-species community responds to different fishing exploitation strategies, particularly: "boats following high-yield boats (Cartesian); boats fishing at random sites (stochast-random); and boats fishing at least exploited sites (stochast-pressure)." They find that the stochast-random strategy is optimal at low fishing pressure, while the Cartesian strategy is more effective at high fishing pressure, in both yield per catch and future biomass growth. Huang and Smith [38] solve a dynamic structural model with strategic interactions for fishing shrimp with two types of externalities. They model a cost due to a stock externality, which has an overall negative impact on utility, as well as a cost due to congestion externality, which turns out to have a positive effect on utility.

Management of international fishing stocks can be challenging because of the difficulty of defining and assigning fishing rights [39, 47]. There are numerous controversies, including "disagreements over the meaning and intent of

[1]"A biological fish stock is a group of fish of the same species that live in the same geographic area and mix enough to breed with each other when mature. A management stock may refer to a biological stock, or a multispecies complex that is managed as a single unit" [1].

fishing rights, disputes over the distribution of rights and associated economic gain, and concern for disruptions imposed" [39] on those operating under the previous system. Many countries use quota systems to manage national fisheries [30, 54], and economists have described methods for envy-free allocation of such quotas [42]. Froese [34] notes that "fishing quotas are decided on the basis of political considerations, largely ignoring the scientific advice, and typically legalizing catches beyond safe levels." Echoing the reasoning for simple treaty forms in Chapter 1, he proposes three simple indicators of fish stock health and justifies each of them as a measure for whether there is overfishing happening. The first indicator is "percentage of mature fish in catch, with 100% as target." Attaining this target allows all fish to spawn at least once and maintain healthy stocks. The second indicator is "percent of specimens with optimum length in catch, with 100% as target." Following this rule prevents growth overfishing, and over time increases the size and value of the fish. The third and final indicator is "percentage of 'mega-spawners' in catch, with 0% as the target," which allows large females to lay more eggs and prevents "subsequent recruitment failure."

The link missing in the literature is between which fish are most in danger of extinction and which fish are under current protection, a connection of biology and economics. Using the idea of increasing action gaps developed in the previous chapter, I attempt to provide an explanation for why international coordination occurs for certain species of fish and not others.

3.3 A Simple Fishery

Many of the most difficult-to-analyze externalities are dynamic in nature. If a dynamic game is Markovian in nature, then inheritability in value functions [58] applies to many of the curvature requirements derived in Chapter 2. Therefore, if an externality only enters in the stage game, and does not affect the dynamics in an opposing manner, the earlier results are clearly extendable. The question remains: what of a purely dynamic externality?

Consider a simple stochastic fishing zone model with two agents. The fishing zone can be in one of two states: damaged, with $s = 0$, or productive, with $s = 1$. When the fishing zone is productive, agents take no strategic action and receive a deterministic payoff of u_1. With probability $(1 - r)$, the zone remains productive, while with probability r, the fishing zone becomes damaged.

When the fishing zone is damaged, each agent i has a choice variable $p_i \in \left[\underline{p}, \bar{p}\right]$. This p_i is part of the transition probability back to the productive state, so a higher choice gives a higher chance of getting out of the damaged state. However, this action is costly in terms of yield. The reward in the bad state is $u_0(p_i)$, a function that is decreasing in p_i and dominated by the good state's utility, i.e. $u_1 > u_0(\underline{p}) > u_0(\bar{p}) > 0$ and $u_0'(p_i) < 0$. The agent's utility is only a function of his own choice, p_i, so there are no static externalities. The transition probability, on the other hand is a function of both agents' choices:

$$P_{0,1} = p_i + p_j$$

To keep $P_{0,1}$ in the range of probability, it must be that $2\bar{p} \leq 1$. Because it is a function of both agents actions, the transition probability displays a positive externality.

The transition matrix is:

$$
\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{s_{t+1}} \\
 & & 0 & 1 \\
\cline{3-4}
\multirow{2}{*}{s_t} & 0 & 1 - p_i - p_j & p_i + p_j \\
\cline{3-4}
 & 1 & r & 1 - r \\
\cline{3-4}
\end{array}
$$

Consider a stationary strategy as a candidate for a Markov perfect equilibrium. The strategy would be to play p_i^M whenever the state is damaged, or bad, satisfying the following value functions:

$$
\begin{aligned}
\Pi_g^i(p_i, p_j^M) &= u_1 + \beta\left[(1-r)\Pi_g^i(p_i, p_j^M) + r\Pi_b^i(p_i, p_j^M)\right] \\
\Pi_b^i(p_i, p_j^M) &= u_0(p_i) + \beta\big[(p_i + p_j^M)\Pi_g^i(p_i, p_j^M) \\
&\quad + (1 - p_i - p_j^M)\Pi_b^i(p_i, p_j^M)\big]
\end{aligned}
\tag{3.1}
$$

Since this is stationary, the good-state value function, $\Pi_g^i(p_i, p_j^M)$, can be solved for directly as a function of the other variables. The derivation is in the Appendix, and it is equal to:

$$
\Pi_g^i(p_i, p_j^M) = \frac{u_1 + \beta r \Pi_b^i(p_i, p_j^M)}{(1 - \beta(1-r))}
$$

This can be substituted back into the bad-state value function, solving for $\Pi_b^i(p_i, p_j^M)$. This is completed in the Appendix, and the resulting function is:

$$\Pi_b^i(p_i, p_j^M) = \frac{(1-\beta(1-r))\,u_0(p_i) + \beta(p_i + p_j^M)u_1}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]}$$

Therefore, agent i's maximization problem is as follows, given player j's Markov perfect stationary action:

$$\begin{aligned} \max_{p_i} \quad & \frac{(1-\beta(1-r))\,u_0(p_i) + \beta(p_i + p_j^M)u_1}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]} \\ \text{s.t.} \quad & p_i \geq \underline{p} \\ & p_i \leq \bar{p} \end{aligned}$$

Lemma 3.1. *There exists a unique interior Markov perfect equilibrium described by:*

$$\begin{aligned} \beta u_0(p_i^M) - \left[1-\beta+\beta r+\beta\left(p_i^M+p_j^M\right)\right] u_0'(p_i^M) - \beta u_1 = 0 \\ \beta u_0(p_j^M) - \left[1-\beta+\beta r+\beta\left(p_i^M+p_j^M\right)\right] u_0'(p_j^M) - \beta u_1 = 0 \end{aligned}$$

The proof is in the Appendix. If the agents are symmetric, then this becomes:

$$\beta u_0(p^M) - \left[1-\beta+\beta r+2\beta p^M\right] u_0'(p^M) - \beta u_1 = 0,$$

which can be rearranged to form:

$$\beta\left(u_0(p^M) - u_1\right) = \left[1-\beta+\beta r+2\beta p^M\right] u_0'(p^M),$$

a statement which has the usual interpretation of equating marginal cost and marginal benefit and where both sides are negative.

Now consider a cooperative version of the stationary Markov equilibrium.[2] The value functions are as follows:

$$\begin{aligned}
\Pi_g^P(p_i, p_j) &= 2u_1 + \beta\left[(1-r)\Pi_g^P(p_i, p_j) + r\Pi_b^P(p_i, p_j)\right] \\
\Pi_b^P(p_i, p_j) &= u_0(p_i) + u_0(p_j) + \beta\big[(p_i + p_j)\Pi_g^P(p_i, p_j) \\
&\qquad + (1 - p_i - p_j)\Pi_b^P(p_i, p_j)\big]
\end{aligned} \tag{3.2}$$

Similarly as for the non-cooperative case, the good-state value function, $\Pi_g^P(p_i, p_j)$, can be solved for explicitly, then substituted into the bad-state function, which is also obtained. The derivations are in the Appendix and give the social planner's maximization problem:

$$\begin{aligned}
\max_{p_i, p_j} \quad & \frac{(1-\beta(1-r))\left(u_0(p_i) + u_0(p_j)\right) + 2\beta(p_i + p_j)u_1}{(1-\beta)\left[1 - \beta + \beta r + \beta\left(p_i + p_j\right)\right]} \\
\text{s.t.} \quad & p_i \geq \underline{p} \\
& p_i \leq \bar{p} \\
& p_j \geq \underline{p} \\
& p_j \leq \bar{p}
\end{aligned}$$

Lemma 3.2. *There exists a unique interior Markov perfect cooperative equilibrium given by:*

$$\beta\left[u_0(p_i) + u_0(p_j)\right] - \left[1 - \beta + \beta r + \beta\left(p_i + p_j\right)\right] u_0'(p_i) - 2\beta u_1 = 0$$

$$\beta\left[u_0(p_i) + u_0(p_j)\right] - \left[1 - \beta + \beta r + \beta\left(p_i + p_j\right)\right] u_0'(p_j) - 2\beta u_1 = 0$$

[2]There are two ways to define the social planner's corresponding restrictions on actions: he could keep each p_i within $[\underline{p}, \bar{p}]$, or could keep the sum $p_i + p_j$ within $[2\underline{p}, 2\bar{p}]$. The second method gives the social planner some extra transferability the players themselves do not have, so the first method corresponds more naturally to the problem.

This is proven in the Appendix. With symmetric agents, have one condition for p^P:

$$2\beta u_0(p^P) - \left[1 - \beta + \beta r + 2\beta p^P\right] u_0'(p^P) - 2\beta u_1 = 0$$

which can be rearranged as:

$$2\beta\left(u_0(p^P) - u_1\right) = \left[1 - \beta + \beta r + 2\beta p^P\right] u_0'(p^P)$$

The simplified symmetric case will be used. A comparison of the symmetric non-cooperative and cooperative first order conditions is below:

$$\beta\left(u_0(p^M) - u_1\right) = \left[1 - \beta + \beta r + 2\beta p^M\right] u_0'(p^M)$$
$$2\beta\left(u_0(p^P) - u_1\right) = \left[1 - \beta + \beta r + 2\beta p^P\right] u_0'(p^P)$$

For continuation purposes, the following assumptions are imposed:

1. Dominance of the good state, i.e.

$$u_1 > 2u_0(\underline{p}) \tag{3.3}$$

With this assumption, the utility in the good state is sufficiently large, larger than twice the maximum possible utility in the bad state. This is to ensure that the good state is tempting enough.

2. Weak concavity, i.e.

$$0 \geq u_0''(p) \tag{3.4}$$

With this assumption, the utility in the bad state is weakly concave in p.

3. Sufficiently high discount factor β, i.e.

$$[1-\beta+\beta r+2\beta p^{P}]>0 \tag{3.5}$$

This assumption will guarantee a few non-zero denominators. What it basically means is that the r given and p^P chosen are sufficiently large together that their discounted sum is larger than $1-\beta$. With exogenous parameters only, the following is sufficient for the condition above:

$$[1-\beta+\beta r]>0 \tag{3.6}$$

The first matter is to demonstrate that there is an externality situation, in that the social planner recommends a higher probability and promotes the positive externality.

Lemma 3.3. *Under the assumptions listed, the social planner's symmetric action p^P is larger than the non-cooperative action p^M.*

The proof is in the Appendix. Having established the externality, the next question is whether any of the parameters in the problem cause an increasing action gap. Hence, the comparative statics of the problem are addressed next.

Lemma 3.4. *Under the assumptions listed, the following are the comparative statics of the problem:*

- *With respect to the discount factor, β, the non-cooperative action is increasing:*

$$\frac{\partial p^M}{\partial \beta}>0,$$

as is the cooperative action:

$$\frac{\partial p^P}{\partial \beta} > 0.$$

- *With respect to the static good-state reward, u_1, the non-cooperative action is increasing:*

  $$\frac{\partial p^M}{\partial u_1} > 0,$$

  as is the cooperative action:

  $$\frac{\partial p^P}{\partial u_1} > 0.$$

- *With respect to the transition from the good state to the bad state, r, the non-cooperative action is decreasing:*

  $$\frac{\partial p^M}{\partial r} < 0,$$

  while the cooperative action is constant:

  $$\frac{\partial p^P}{\partial r} = 0.$$

The proof deriving these comparative statics is in the Appendix. The last two derivatives, those with respect to r, point the direction of investigation of a parameter making the externality worse.

The variable r is the probability of transitioning to the bad state if in the good state. This number is technically not part of the externality story at all, since it is unaffected by the players' actions and is simply a "fact of life."

The social planner pays it no mind in determining the optimal action for the players; as can be seen, if r increases, i.e. the likelihood of staying in the good state decreases, the social planner does not change the action taken, $\frac{\partial p^P}{\partial r} = 0$.

If $\beta u_0'(p^M) > \left(1 - \beta + \beta r + 2\beta p^M\right) u_0''(p^M)$, then the non-cooperative action is shrinking because less of the costly action is taken to get back to good state. With higher probability of leaving the good state for the bad, the good state is less valuable than before, so sacrificing utility to return to it is not desirable. This reaction, however, creates a pattern which worsens the dynamic externality and provides an example of an increasing action gap.

3.4 A Realistic Fishery

The previous fishery model simplified reality, with a focus on understanding how a transition probability can cause a dynamic externality, even when a one-stage game does not contain a commons problem. This section presents a fishery model which is slightly more realistic. The primitives of this model are:

1. agents, $I = \{1, 2\}$,
2. actions, $A_i = [0, 1]$,
3. utility $u_i(x_i)$, where x_i is a function of $a \in A \equiv \times_{i \in I} A_i$, and
4. discount factor, δ.

Agents come into period knowing residual stock, y_{t-1}, and last period's shock, z_{t-1}. There is some growth shock, z_t, which is applied to the residual stock, but this shock is not revealed to the players. There is perhaps some knowledge about the expectation of z_t from z_{t-1}, depending the distribution of the growth shocks and any knowledge contained therein. When the growth function is applied, $\hat{y}_t$ is the new starting stock.

Agent i takes action a_{it}, while agent j takes action a_{jt}, each to maximize his own expected value. This action results in the individual catch, x_{it}, according to deterministic harvest function. The individual catches are observed by both players, who can then calculate the total summed catch. From this information, players can then back out the grown stock, as well as the period's growth shock. Players then know the new residual stock, which is the grown stock less the catches.

The particular functional forms are now described. First, the period utility, $u(x_{it})$, should be something relatively simple, given that the growth process will be more complicated. The utility function should be increasing in own catch and possibly weakly concave, depending on desired risk preferences. Notably, the function does not take any other arguments, so the agent is not harmed by the fact that the other agent may have caught something as well, nor does the agent receive any direct benefit from the stock's existence. The simplest possible function to use here is a linear one:[3]

$$u(x_{it}) = x_{it}. \tag{3.7}$$

The catch x_{it} is determined by the harvest function, $h(a_{it}, a_{jt}, \hat{y}_t)$. This is a deterministic function, though the grown stock, $\hat{y}_t$, is unknown to the players at the time of their decisions. The harvest function should be increasing and concave in own action, a_{it}, as more effort increases catch but with diminishing marginal returns. On the other hand, the function should be decreasing in the other player's action, a_{jt}, as more effort on an opponent's part decreases a player's own catch. This aspect captures a direct negative externality in the stage game, a characteristic which was not present in the example in Section 3.3. Furthermore, the harvest function should be submodular, so more effort on an opponent's part decreases the marginal catch as well. Finally, the harvest function should be increasing in stock. A possible function to use here is:

$$h(a_{it}, a_{jt}, \hat{y}_t) = \hat{y}_t \left(1 - \exp\left(-a_{it}\left(s - a_{jt}\right)\right)\right). \tag{3.8}$$

It is assumed that s is greater than one, to make sure that $s - a_{jt}$ is a positive amount.

Working backwards, the next function to characterize is the growth function. This function needs to correspond to the biological principles of fish populations. The growth function should be increasing in the residual stock

[3] A logarithmic utility function would also be a good choice.

and in the shock. There is also a natural growth rate, r, which is generally somewhere between $(0, 1)$, which enters the growth function in an increasing manner. A possible function to serve as the growth function is:

$$g(y_{t-1}, z_t) = y_t \exp(r(1 + z_t)). \tag{3.9}$$

The growth shock, z_t, is modeled in a manner that allows for possible correlation over time. One possibility is an AR1 process with some idiosyncratic shock. The growth shock can be written in the following manner:

$$z_t = \rho z_{t-1} + \varepsilon_t, \tag{3.10}$$

where ρ is the coefficient of correlation and ε_t is distributed according to a standard normal. If ρ is zero, then there is no time correlation, and z_t is identical to the idiosyncratic shock. As ρ increases, the time correlation of growth shocks increases.

Using these functions and the set-up, the maximization problem of agent i, given a_{jt}, can then be written as:

$$\begin{aligned} V_i(y_{t-1}, z_{t-1}) = \max_{a_i \in [0,1]} \ & E\Big[y_{t-1} \exp(r(1 + \rho z_{t-1} + \varepsilon_t)) \\ & \cdot (1 - \exp(-a_{it}(s - a_{jt}))) \\ & + \delta V_i\Big(y_{t-1} \exp(r(1 + z_t))(-1 + \exp(-a_{it}(s - a_{jt}) \\ & + \exp(-a_{jt}(s - a_{it}))), 1 + \rho z_{t-1} + \varepsilon_t \Big) \Big| \varepsilon_t \Big] \\ \text{s.t. } \varepsilon_t \sim N(0,1), \ \ & f(\varepsilon_t) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2} \right) \end{aligned} \tag{3.11}$$

The exact bounds are not addressed here, though there must be some knowledge of what occurs if the stock goes extinct. This is an extension to be covered when estimation is attempted.

In order to apply the parameterization framework in a negative externality, the first step is to characterize the nature of the externality and parameters that may affect it. There were three possible linear effects describe: the opponent effect, the own effect, and the submodular effect.

Recall that the opponent effect involved directly increasing the negative externality caused by the opponent's action, the own effect increases the marginal benefit of a player's own action, and the submodular effect decreases an already-negative cross-partial. Therefore, in order to determine which parameterization to use, the following derivatives should be taken:

1. If a parameter θ causes an own effect, then it should increase the marginal utility of own action. Therefore, the following derivative should be taken and checked:
$$\frac{\partial}{\partial \theta}\left[\frac{\partial u_i}{\partial a_i}\right] > 0$$
and/or
$$\frac{\partial}{\partial \theta}\left[\frac{\partial V_i}{\partial a_i}\right] > 0.$$
2. If a parameter θ causes an opponent effect, then it should make the negative externality stronger. Therefore, the following derivative should be taken and checked:
$$\frac{\partial}{\partial \theta}\left[\frac{\partial u_i}{\partial a_j}\right] < 0$$

and/or

$$\frac{\partial}{\partial \theta}\left[\frac{\partial V_i}{\partial a_j}\right] < 0.$$

3. If a parameter θ causes a submodular effect, then it should make a negative cross-partial even more negative. Therefore, the following derivative should be taken and checked:

$$\frac{\partial}{\partial \theta}\left[\frac{\partial^2 u_i}{\partial a_i \partial a_j}\right] < 0$$

and/or

$$\frac{\partial}{\partial \theta}\left[\frac{\partial^2 V_i}{\partial a_i \partial a_j}\right] < 0.$$

For a set-up with a positive externality or supermodularity, the signs of interest would be opposite.

The second step is to determine if the necessary curvatures for Theorems 2.1 and 2.2 to hold are met. For the derivatives that hold true, the corresponding linearized parameters can be added to the baseline utility function, and with the proper curvature, they give the directions described in the previous chapter. From the results derived, two sufficient situations are the most clear: if the opponent effect is the only effect, then the action gap holds by Theorem 2.1; if the own effect is the only effect, then the requirements for Theorem 2.2 must be checked.

Checking the derivatives of a dynamic function can be a difficult task, particularly with a nuanced problem like Equation (3.11). Though the full derivative should be checked, the following analysis will focus on just the first stage. The linearization framework could be extended to two parameters, one which describes the effect of the parameter on the first stage, and one which describes the effect on continuation value. This is an extension which may prove interesting, but at this moment, only the stage parameter is investigated.

If just examining the first stage, then the maximization problem becomes:

$$\begin{aligned} U \equiv \max_{a_{it}} \int_{-\infty}^{\infty} & y_{t-1} \exp(r(1+\rho z_{t-1}+\varepsilon_t))[1-\exp(-a_{it}(s-a_{jt}))] \\ & \cdot \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t \end{aligned} \tag{3.12}$$

The parameter of interest is ρ, the coefficient of correlation on the growth shock. With zero correlation, there is no information gained from knowing the previous period's shock. With positive correlation, the previous period's growth shock gives some information, however, and may affect the strategies used and the negative externality.

To check if there is an own effect, the first step is to take the derivative of U with respect to a_i:

$$\begin{aligned} \frac{\partial U}{\partial a_i} = & \int_{-\infty}^{\infty} y_{t-1} \exp(r(1+\rho z_{t-1}+\varepsilon_t))[-\exp(-a_{it}(s-a_{jt}))(-1)(s-a_{jt})] \\ & \cdot \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t \\ = & \int_{-\infty}^{\infty} y_{t-1} \exp(r(1+\rho z_{t-1}+\varepsilon_t)) \exp(-a_{it}(s-a_{jt}))(s-a_{jt}) \\ & \cdot \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t \end{aligned} \tag{3.13}$$

This derivative is positive, as the larger action increases one-stage utility. However, since today's action lowers tomorrow's stock, this would enter negatively into continuation value, making the derivative with respect to the full value function have an ambiguous sign.

The next step is to look at the effect of ρ on this derivative:

$$
\begin{aligned}
\frac{\partial}{\partial \rho}\left[\frac{\partial U}{\partial a_i}\right] = & \int_{-\infty}^{\infty} y_{t-1} \exp(r(1+\rho z_{t-1}+\varepsilon_t)) r z_{t-1} \\
& \cdot \exp(-a_{it}(s-a_{jt}))(s-a_{jt}) \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t \\
& = y_{t-1} z_{t-1} \exp(-a_{it}(s-a_{jt}))(s-a_{jt}) \\
& \cdot \int_{-\infty}^{\infty} r \exp(r(1+\rho z_{t-1}+\varepsilon_t)) \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t \\
& = y_{t-1} z_{t-1} \exp(-a_{it}(s-a_{jt}))(s-a_{jt}) \\
& \cdot \int_{-\infty}^{\infty} r \exp(r(1+\rho z_{t-1})) \exp(r\varepsilon_t) \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t \\
& = y_{t-1} z_{t-1} \exp(r(1+\rho z_{t-1})) \exp(-a_{it}(s-a_{jt}))(s-a_{jt}) \\
& \cdot \int_{-\infty}^{\infty} \frac{r}{\sqrt{2\pi}} \exp(r\varepsilon_t) \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t
\end{aligned}
\tag{3.14}
$$

To determine the sign of this expression, each piece must be examined. With proper bounds on extinction, then y_{t-1} is positive, as is $s - a_{jt}$ and each of the exponential functions, including the integrated term. However, the sign of z_{t-1} is unknown. Because this growth shock is hit by an idiosyncratic shock distributed according to a standard normal, it is possible for the growth shock to be positive or negative. The entirety of the expression is positive when previous period's growth shock, z_{t-1} is positive, but it is negative if that shock is negative. Since this value could foreseeably switch every period, not much information is gained here as to whether an own effect is present.

To check if there is an opponent effect, the first step is to take the derivative of U with respect to a_j:

$$\begin{aligned}\frac{\partial U}{\partial a_j} =& \int_{-\infty}^{\infty} y_{t-1}\exp(r(1+\rho z_{t-1}+\varepsilon_t)) \\ &\cdot [-\exp(-a_{it}(s-a_{jt}))(-a_{it})(-1)]\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-\varepsilon_t^2}{2}\right)\mathrm{d}\varepsilon_t \\ =& -\int_{-\infty}^{\infty} y_{t-1}\exp(r(1+\rho z_{t-1}+\varepsilon_t)) \\ &\cdot \exp(-a_{it}(s-a_{jt}))a_{it}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-\varepsilon_t^2}{2}\right)\mathrm{d}\varepsilon_t\end{aligned} \tag{3.15}$$

This derivative is negative, as a larger action taken by the opponent decreases one-stage utility. The opponent's action also lowers tomorrow's stock, so this enters negatively into the continuation value. Unlike the own action, the opponent action enters into both parts in the same manner, and so the derivative with respect to the full value function should be negative.

The next step is to determine the effect of ρ on this derivative:

$$
\begin{aligned}
\frac{\partial}{\partial \rho}\left[\frac{\partial U}{\partial a_j}\right] =& \int_{-\infty}^{\infty} y_{t-1}\exp(r(1+\rho z_{t-1}+\varepsilon_t))rz_{t-1} \\
&\cdot \exp(-a_{it}(s-a_{jt}))(-a_{it})\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-\varepsilon_t^2}{2}\right)\mathrm{d}\varepsilon_t \\
=& \, y_{t-1}z_{t-1}\exp(-a_{it}(s-a_{jt}))(-a_{it}) \\
&\cdot \int_{-\infty}^{\infty} r\exp(r(1+\rho z_{t-1}+\varepsilon_t))\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-\varepsilon_t^2}{2}\right)\mathrm{d}\varepsilon_t \\
=& -y_{t-1}z_{t-1}\exp(r(1+\rho z_{t-1}))\exp(-a_{it}(s-a_{jt}))a_{it} \\
&\cdot \int_{-\infty}^{\infty} \frac{r}{\sqrt{2\pi}}\exp(r\varepsilon_t))\exp\left(\frac{-\varepsilon_t^2}{2}\right)\mathrm{d}\varepsilon_t
\end{aligned}
\tag{3.16}
$$

Again, most of the pieces of this function are positive, but there is a negative sign out front. Like with the own effect, the sign is determined by the previous period's growth shock, z_{t-1}. If that is positive, then the expression is negative, while if that growth shock is negative, then the whole expression is positive.

To check if there is a submodular effect, the first step is to take the derivative of U with respect to a_i and a_j:

$$
\begin{aligned}
\frac{\partial^2 U}{\partial a_i \partial a_j} = & \int_{-\infty}^{\infty} y_{t-1} \exp(r(1+\rho z_{t-1}+\varepsilon_t))[\exp(-a_{it}(s-a_{jt}))(-1) \\
& + \exp(-a_{it}(s-a_{jt}))(-1)(s-a_{jt})(-a_{it})] \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t \\
= & \int_{-\infty}^{\infty} y_{t-1} \exp(r(1+\rho z_{t-1}+\varepsilon_t))[-\exp(-a_{it}(s-a_{jt})) \\
& - \exp(-a_{it}(s-a_{jt}))(-a_{it}(s-a_{jt}))] \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t \\
= & -\int_{-\infty}^{\infty} y_{t-1} \exp(r(1+\rho z_{t-1}+\varepsilon_t)) \exp(-a_{it}(s-a_{jt})) \\
& \cdot (1-a_{it}(s-a_{jt})) \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t
\end{aligned} \tag{3.17}
$$

This derivative is weakly negative, since $s - a_{jt} > 0$, while $a_{it} \leq 1$ because of bounds on the domain. Hence there is weak submodularity displayed in the problem.

The next step is to determine the effect of ρ on this cross-partial:

$$
\begin{aligned}
\frac{\partial}{\partial \rho}\left[\frac{\partial^2 U}{\partial a_i \partial a_j}\right] = & -\int_{-\infty}^{\infty} y_{t-1} \exp(r(1+\rho z_{t-1}+\varepsilon_t)) r z_{t-1} \exp(-a_{it}(s-a_{jt})) \\
& \cdot (1-a_{it}(s-a_{jt})) \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t \\
= & -y_{t-1} z_{t-1} \exp(-a_{it}(s-a_{jt}))(1-a_{it}(s-a_{jt})) \\
& \cdot \int_{-\infty}^{\infty} r \exp(r(1+\rho z_{t-1})) \exp(r\varepsilon_t)) \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t \\
= & -y_{t-1} z_{t-1} \exp(r(1+\rho z_{t-1})) \exp(-a_{it}(s-a_{jt})) \\
& \cdot (1-a_{it}(s-a_{jt})) \int_{-\infty}^{\infty} \frac{r}{\sqrt{2\pi}} \exp(r\varepsilon_t)) \exp\left(\frac{-\varepsilon_t^2}{2}\right) \mathrm{d}\varepsilon_t
\end{aligned} \tag{3.18}
$$

As with the opponent effect, most pieces of this expression are positive, there is a negative sign out front, and the actual sign of the expression depends on the previous period's growth shock, z_{t-1}. If that shock is positive, then the whole expression is negative, while if that shock is negative, then the expression is positive.

3.5 Conclusions

These two chapters work together to provide the very beginning of insight into a complicated problem of externalities. With the number of possibilities for coordination to reduce an externality problem, there must be some method to determine which situations merit that coordination. I put forth the framework examining externalities based on some parameter which causes increasing action gaps between coordination and non-coordination. Since many externality stories can be difficult to analyze, I proposed a method of linearization based on three possible effects and analyzed two notable cases of increasing action gaps. The first was a sole opponent effect in a symmetric game, where the optimal action under coordination unambiguously diverges from the non-coordination action. The second was an own effect in a potentially non-symmetric game, where sufficient conditions for divergence include accelerating benefits to reduction of the action that causes the negative externality.

The main extension to pursue is that of centered parameterizations. Unlike the simple linear parameterization, a centered term can give better insight into the utility gap as well. However, there must be careful understanding of how the separate centerings affect economic intuition. In early analysis, the centered Taylor expansion form suggests that the submodular effect is null. In comparison to the ambiguous decreases under the linear parameterization, this departure suggests that the framework should be checked for robustness to parameterization.

The difficulties in applying the framework in this chapter suggest an even more complicated problem is at hand than originally thought. Further work requires more conclusive examples than the fishery models presented here. For instance, a modification of the full fishery model which guarantees that growth shocks are between zero and two would be less realistic, perhaps, but would guarantee the necessary sign for each of the three effects to be included. However, the initial analysis presented here does suggest that persistence in growth shocks of fish gives some sort of increase in action gap, and thereby a motivation for coordination under high time correlation.

Appendices

Appendix A

Proofs and Derivations for Chapter 1

A.1 General Set-up and Unrestricted Commitment Power

Proof of Lemma 1.1.

Restatement of Lemma 1.1. *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 positive ε, $\varepsilon^2 < \varepsilon$.

□

Proof of Lemma 1.2.

Restatement of Lemma 1.2. *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 the actions in summed utility.*

Proof. By Lemma 1.1, $a^{Un}(u, I)$ 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. □

A.2 Lump-sum Commitment Power

A.2.1 Proofs for Section 1.4.1

Proof of Theorem 1.1.

Restatement of Theorem 1.1. *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. For some I, 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 for which exclusion is optimal and all the assumptions on $\mathcal{U}$ hold, and then demonstrate that those assumptions are open conditions.

The particular function considered is

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

The parameter θ is as described in Section 1.4.2, where the group J consists of agents who have $\theta_i = 1$, while the remaining agents not in J have $\theta_i = \theta \in \Theta = (0, 1)$.

Exclusion Result

An example of the optimality of exclusion under sufficient heterogeneity for this chosen function comes from Lemma 1.4, which is itself proven in Appendix A.2.2.

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

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

  - Second derivatives

$$\frac{d^2u_i(a)}{da^2} = \frac{\partial^2 u_i(a)}{\partial a_i^2} + 2\sum_{j\neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j} + \sum_{j\neq i} \left\{ \frac{\partial^2 u_i(a)}{\partial a_j^2} + \sum_{k\neq i \,\mathrm{or}\, j} \frac{\partial u_i(a)}{\partial a_j \partial a_k} \right\}$$

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

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

$$\forall\, j \neq i \;\; \frac{\partial u_i(a)}{\partial a_j^2} = 0$$

$$\forall\, k \neq i \,\mathrm{or}\, j \;\; \frac{\partial u_i(a)}{\partial a_j \partial a_k} = 0$$

- *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 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 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.

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

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

  The Lagrangian is:

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

  The Kuhn-Tucker conditions are:

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

  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:

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

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 $\pm$):

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

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

This means that a_i^{*-} is not within the domain, and so it cannot be an 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:

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

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:

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

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

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

Subtract the bottom condition from the top one:

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

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.

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

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 difference 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 leads 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.

- *Unique Nash equilibrium:* For the purposes of Theorem 1.1, it is sufficient for the property of unique Nash equilibrium to be open conditional upon the previous properties.

Let $\mathcal{U}_{conc}$ denote the $C^2(A)$ utility functions which are strictly concave in own actions. Observe that our functions of interest are within this set, $u \in \mathcal{U} \subset \mathcal{U}_{conc}$. Let $Eq(u)$ denote the set of equilibria for the game $\Gamma = (u_i, A_i)_{i \in I}$. The vector of first order conditions can then be denoted $F(u, a) \in \mathbb{R}^I$ and is written as:

$$F(u, a) = \begin{pmatrix} \frac{\partial u_1(a)}{\partial a_1} \\ \frac{\partial u_2(a)}{\partial a_2} \\ \vdots \\ \frac{\partial u_n(a)}{\partial a_n} \end{pmatrix}$$

Let A^o denote the interior of A. By strict own concavity, we know that for an action profile in the interior, $a^* \in A^o$, we have that it is an equilibrium, $a^* \in Eq(u)$, if and only if the FOC vector is equal to zero at that profile, $F(u, a^*) = 0$.

Thinking of openness of the condition of unique Nash equilibrium, we need that for any small vector movement in others' actions, agent i has one unique best action close to his previous action. Running into boundaries and corner solutions might initially be of concern, so we will carefully consider the implicit function theorem and the determinant of the FOC.

For a non-empty $J \subset I$, let $DF_J(u, a)$ denote the determinant of the FOCs for J. Let the cardinality of J be equal to m, so $DF_J(u, a)$ is an $m \times m$ matrix which can be written as:

$$D_J F(u,a) = \begin{pmatrix} \frac{\partial^2 u_1(a)}{\partial u_1^2} & \frac{\partial^2 u_1(a)}{\partial a_1 \partial a_2} & \cdots & \frac{\partial^2 u_1(a)}{\partial a_1 \partial a_m} \\ \frac{\partial^2 u_2(a)}{\partial a_1 \partial a_2} & \frac{\partial^2 u_2(a)}{\partial a_2^2} & \cdots & \frac{\partial^2 u_2(a)}{\partial a_2 \partial a_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 u_m(a)}{\partial a_1 \partial a_m} & \frac{\partial^2 u_m(a)}{\partial a_2 \partial a_m} & \cdots & \frac{\partial^2 u_m(a)}{\partial a_m^2} \end{pmatrix}$$

The next result gives the invertibility of the FOC at the equilbrium, $DF_I(u, a^*)$, for these concave functions.

Lemma A.1. *The set of $u \in \mathfrak{U}_{conc}$ such that the J-determinant is invertible at any equilibrium, i.e. $DF_J(u, a^*) \neq 0$ for all $a^* \in Eq(u)$ and all non-empty $J \subset I$, is a non-empty open subset of $\mathfrak{U}$.*

Proof. Non-emptiness can be shown by example.

Denote two sets of interest in $\mathcal{U}_{conc}$. Let B_1 be the set of utility functions and their equilibria in the concave functions, denoted formally as $B_1 = \{(u, a) \in \mathcal{U}_{conc} \times A : a \in Eq(u)\}$. Let $B_{2,J}$ be the set of utility function and actions for which the J-determinant is equal to zero, formally denoted as $B_{2,J} = \{(u, a) \in \mathcal{U}_{conc} \times A : DF_J(u, a) = 0\}$. Finally, let B denote the projection of of $B_1 \cap (\cup_{J \subset I} B_{2,J})$ onto $\mathcal{U}_{conc}$.

Both B_1 and $\cup_{J \subset I} B_{2,J}$ are closed, so their intersection is as well. Because A is compact, the projection of a closed subset of $\mathcal{U}_{conc} \times A$ onto $\mathcal{U}_{conc}$ is closed.The complement of the closed set B is the requisite set, with invertible J-determinants at equilibrium, and it is open. $\square$

Let $\mathcal{U}_{inv}$ denote the set of invertible u from Lemma A.1. For $u \in \mathcal{U}_{inv}$ and $a^* \in Eq(u)$, we say that (u, a^*) is **not flat at the boundary** if either the equilibrium is in the interior of A, i.e. $a^* \in A^o$, or if for each $i \in I$ with equilibrium action a_i^* in the boundary of A_i, the gradient of u_i points outwards. For instance, if $a_i^* = 0$, then $\frac{\partial u_i(a^*)}{\partial a_i} < 0$, so the agent would want to decrease more if he could, or if $a_i^* = 1$, then $\frac{\partial u_i(a^*)}{\partial a_i} > 0$, so the agent would want to increase more if they could. This notion sets us up for the next lemma and the invertibility result.

Lemma A.2. *There exists a non-empty, open set of $u \in \mathcal{U}_{conc}$ for which there is a unique equilibrium, $\#Eq(u) = 1$, and for which $a^*(u)$ is a smooth function of u.*

Proof. Let C denote the set of $u \in \mathcal{U}_{inv}$ for which there is an $a^* \in Eq(u)$ that is flat at the boundary, which means that every $a^* \in Eq(u)$ has the property that for all $i \in I$, a_i^* is in the boundary of A_i. The set C is closed.

Let $\mathcal{U}'_{inv}$ denote the complement of C in $\mathcal{U}_{inv}$. Suppose that $u \in \mathcal{U}'_{inv}$ has one equilibrium and that $J \subset I$ is the non-empty subset of I for which a_j^* is not in the boundary of A_j. Since $u \in \mathcal{U}_{inv}$, then $DF_J(u, a^*) \neq 0$. All of this means that $a^*(\cdot)$ is a locally unique, differentiable function on a neighborhood of u when we hold a_i^* fixed, $i \notin J$. Since (u, a^*) is not flat on the boundary, for small enough changes in u, leaving a_i^* fixed is still optimal for $i \notin J$. Since each u with one equilibrium has such a neighborhood, the set of u with one equilibrium is open. □

This shows that the property of unique Nash equilibrium is open.

Each of the conditions has been shown to be open individually, either unconditionally or conditionally upon the remaining conditions. 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.2.2 Proofs for Section 1.4.2

Proof of Lemma 1.3.

Restatement of Lemma 1.3. *When the group parameter is strictly smaller than one, $\theta < 1$, then the equilibrium action of the players not in J, $a^*_{I\setminus J}(\theta)$, is smaller than the equilibrium action of the players in J, $a^*_J(\theta)$.*

Proof. Subtract FOC of J from FOC of $I \setminus J$.

$$\left[\theta B'\left(a^*_{I\setminus J}(\theta)\right) - B'\left(a^*_J(\theta)\right)\right] = \left[a^*_{I\setminus J}(\theta) - a^*_\theta(\theta)\right] c'\left(\sum_{k\in I\setminus J} a^*_{I\setminus J}(\theta) + \sum_{j\in J} a^*_J(\theta)\right)$$

Suppose not.

Case i. Suppose that when $\theta < 1$, $a^*_{I\setminus J}(\theta) = a^*_J(\theta)$. Then RHS $= 0 \Rightarrow$ LHS should be zero as well. However, LHS < 0. Contradiction shown.

Case ii. Suppose that when $\theta < 1$, $a^*_{I\setminus J}(\theta) > a^*_J(\theta)$. Then RHS $> 0 \Rightarrow$ LHS should be greater than zero as well. However,

$$B''(\cdot) < 0 \Rightarrow \left[\left(a^*_{I\setminus J}(\theta) > a^*_J(\theta)\right) \Rightarrow \left(B'\left(a^*_{I\setminus J}(\theta)\right) < B'\left(a^*_J(\theta)\right)\right)\right].$$

Since $0 < \theta < 1$, then $\theta B'\left(a^*_{I\setminus J}(\theta)\right) < B'\left(a^*_J(\theta)\right)$, which gives LHS < 0. Contradiction shown.

Hence, it must be the case that $a^*_{I\setminus J}(\theta) < a^*_J(\theta)$ when $\theta < 1$. $\square$

<u>Kuhn-Tucker conditions for lump-sum reduction by coalition of the whole</u>

Using the cardinalities defined earlier, the maximization problem can be rewritten as:

$$\begin{aligned}\max_{r\in[0,a^*_{I\setminus J}(\theta)]} \quad & m\Big[B(a^*_J(\theta)-r)-(a^*_J(\theta)-r)\\ & \cdot c\Big(m(a^*_J(\theta)-r)+(n-m)(a^*_{I\setminus J}(\theta)-r)\Big)\Big]\\ & +(n-m)\Big[\theta B(a^*_{I\setminus J}(\theta)-r)-(a^*_{I\setminus J}(\theta)-r)\\ & \cdot c\Big(m(a^*_J(\theta)-r)+(n-m)(a^*_{I\setminus J}(\theta)-r)\Big)\Big]\end{aligned}$$

The Lagrangian is:

$$\begin{aligned}\mathcal{L}(r,\lambda_1,\lambda_2)= \; & m\Big[B(a^*_J(\theta)-r)-(a^*_J(\theta)-r)\\ & \cdot c\Big(m(a^*_J(\theta)-r)+(n-m)(a^*_{I\setminus J}(\theta)-r)\Big)\Big]\\ & +(n-m)\Big[\theta B(a^*_{I\setminus J}(\theta)-r)-(a^*_{I\setminus J}(\theta)-r)\\ & \cdot c\Big(m(a^*_J(\theta)-r)+(n-m)(a^*_{I\setminus J}(\theta)-r)\Big)\Big]\\ & +\lambda_1 r+\lambda_2(a^*_{I\setminus J}(\theta)-r)\end{aligned}$$

The Kuhn-Tucker conditions are:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial r} =& - m\Big[B'(a^*_J(\theta)-r) \\
&-\Big[c\Big(m(a^*_J(\theta)-r)+(n-m)(a^*_{I\setminus J}(\theta)-r)\Big) \\
&+n(a^*_J(\theta)-r)c'\Big(m(a^*_J(\theta)-r)+(n-m)(a^*_{I\setminus J}(\theta)-r)\Big)\Big]\Big] \\
&-(n-m)\Big[\theta B'(a^*_{I\setminus J}(\theta)-r) \\
&-\Big[c\Big(m(a^*_J(\theta)-r)+(n-m)(a^*_{I\setminus J}(\theta)-r)\Big) \\
&+n(a^*_{I\setminus J}(\theta)-r)c'\Big(m(a^*_J(\theta)-r)+(n-m)(a^*_{I\setminus J}(\theta)-r)\Big)\Big]\Big] \\
&+\lambda_1-\lambda_2=0 \\
\lambda_1 r =& 0, \quad \lambda_1 \geq 0, \quad r \geq 0 \\
\lambda_2(a^*_{I\setminus J}(\theta)-r) =& 0, \quad \lambda_2 \geq 0, \quad r \leq a^*_{I\setminus J}(\theta)
\end{aligned}$$

Kuhn-Tucker conditions for lump-sum reduction by J-coalition

$$\max_{\hat{r}\in[0,a^*_J(\theta)]} \; m\left[B(a^*_J(\theta)-\hat{r})-(a^*_J(\theta)-\hat{r})c\Big(m(a^*_J(\theta)-\hat{r})+(n-m)a^J_{I\setminus J}(\theta)\Big)\right]$$

The Lagrangian is:

$$\begin{aligned}
\mathcal{L}(r,\lambda_1,\lambda_2) = \; & m\Big[B(a^*_J(\theta)-\hat{r})-(a^*_J(\theta)-\hat{r}) \\
& \cdot c\Big(m(a^*_J(\theta)-\hat{r})+(n-m)a^J_{I\setminus J}(\theta)\Big)\Big] \\
& +\lambda_1\hat{r}+\lambda_2(a^*_J(\theta)-\hat{r})
\end{aligned}$$

The Kuhn-Tucker conditions are:

$$\frac{\partial \mathcal{L}}{\partial \hat{r}} = - m\Big[B'(a_J^*(\theta) - \hat{r}) - \Big[c\Big(m(a_J^*(\theta) - \hat{r}) + (n-m)a_{I\setminus J}^J(\theta)\Big)$$

$$+m(a_J^*(\theta) - \hat{r})c'\Big(m(a_J^*(\theta) - \hat{r}) + (n-m)a_{I\setminus J}^J(\theta)\Big)\Big]\Big]$$

$$+\lambda_1 - \lambda_2 = 0$$

$$\lambda_1 \hat{r} = 0, \quad \lambda_1 \geq 0, \quad \hat{r} \geq 0$$

$$\lambda_2(a_J^*(\theta) - \hat{r}) = 0, \quad \lambda_2 \geq 0, \quad \hat{r} \leq a_J^*(\theta)$$

Incentive for J-coalition to Reduce.

Let the utility of coalition J taking a lump-sum reduction of r be denoted as

$$u_J^{LS}(r,\theta) = m\left[B(a_J^*(\theta) - r) - (a_J^*(\theta) - r)\, c\left(m(a_J^*(\theta) - r) + (n-m)a_{I\setminus J}^J(\theta)\right)\right].$$

Lemma A.3. *When moving from zero, the players in J in the game in Section 1.4.2 have a strict incentive to increase r and reduce. Formally, this means that $\left.\frac{\partial u_J^{LS}(r,\theta)}{\partial r}\right|_{r=0}$, derived below, is positive:*

$$\left.\frac{\partial u_J^{LS}(r,\theta)}{\partial r}\right|_{r=0} = -m\Big[B'\left(a_J^*(\theta)\right) - c\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^J(\theta)\right)$$

$$- ma_J^*(\theta)c'\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^J(\theta)\right)\Big]$$

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

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

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

At $r = 0$, those in J are agreeing to play the no-coalition Nash equilibrium, meaning that the best responses of those not in J are also the coalition actions, giving $a^*_{I\setminus J}(\theta) = a^J_{I\setminus J}(\theta)$. The first order condition and the derivative of u^{LS}_J 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:

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

Combined with the negative sign on the outside of the parentheses, $\left.\frac{\partial u^{LS}_J(r,\theta)}{\partial r}\right|_{r=0} > 0$. □

Proof of Lemma 1.4.

Restatement of Lemma 1.4. *There exists a threshold value $\bar{\theta} > 0$ for the group parameter such that for all values of the parameter higher than the threshold, $\theta \in (0, \bar{\theta})$, the equilibrium action of players not in J, $a^*_{I\setminus J}(\theta)$, is a binding constraint on problem (1.14).*

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 Φ :

$X \to \mathcal{K}(Y)$ has a closed graph and Y is compact, then Φ is upper hemi-continuous (Corollary 6.1.33 in [26]).

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:

$$f(\sigma, \theta; a_i^o) \equiv \int_A u_i(a;\theta) d\sigma^*(a) - \int_A u_i(a \setminus a_i^o) d\sigma^*(a) \geq 0 \quad \text{(A.2)}$$

  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 (1.10) and (1.11) 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 zero and the equilibrium action $a^*_{I\setminus J}(1) = 0$, then $a^*_{I\setminus J}(\theta)$ for θ arbitarily close to zero is also arbitrarily close to zero.

3. *Binding:* The third step is to show that when solving the coalition I problem for θ close to zero, $a^*_{I\setminus J}(\theta)$ is a binding constraint on choosing $r^*_I(\theta)$. Looking at Equation 1.14, it balances the utility of both groups of players. The players not in 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 A.3. Thus, the action of the players not in 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 not in 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 not in J, giving a Pareto improvement upon the coalition of the whole for some values of the parameter.

$\square$

A.3 Proportional Commitment Power

A.3.1 Proofs for Section 1.4.3

Proof of Theorem 1.2

Restatement of Theorem 1.2. *For any $J \subsetneq I$, $\#J \geq 2$, there is a set of $u \in \mathfrak{U}$ having non-empty interior, for which the vector of actions $a^{Pr}(u, J)$ is conceivable, formally denoted as:*

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

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

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

Proof. For this proof, I use the same approach as for Theorem 1.1, as well as the same function, Equation (A.1). A reminder:

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

The properties of $\mathcal{U}$ have already been shown to be fulfilled, as has their openness. The only thing to prove is that the exclusion result holds for proportional reduction, which is shown through Lemma 1.6, which is itself proven in Appendix A.3.2. The exclusion example in that lemma requires that $B'(0) > 0$ and that the total cost of zero action is also zero. The function used has $B'(a) = \theta$ for any action and has cost of zero when $a_i = 0$.

$\square$

A.3.2 Proofs for Section 1.4.4

Kuhn-Tucker conditions for proportional reduction by coalition of the whole

Using the cardinalities defined earlier, the maximization problem can be rewritten as:

$$\max_{s\in[0,1]} \; m\left[B(sa_J^*(\theta)) - sa_J^*(\theta)c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\right]$$
$$+(n-m)\left[\theta B(sa_{I\setminus J}^*(\theta)) - sa_{I\setminus J}^*(\theta)c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\right]$$

The Lagrangian is:

$$\begin{aligned}\mathcal{L}(s,\lambda_1,\lambda_2) = \; & m\left[B(sa_J^*(\theta)) - sa_J^*(\theta)c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\right] \\ & +(n-m)\Big[\theta B(sa_{I\setminus J}^*(\theta)) - sa_{I\setminus J}^*(\theta)^2 \\ & \quad \cdot c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\Big] \\ & +\lambda_1 s + \lambda_2(1-s)\end{aligned}$$

The Kuhn-Tucker conditions are:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial s} =& m\Bigg[B'(sa^*_J(\theta))a^*_J(\theta) - \bigg[a^*_J(\theta)c\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big) \\
&+ sa^*_J(\theta)c'\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big) \\
&\cdot \big(ma^*_J(\theta) + (n-m)a^*_{I\setminus J}(\theta)\big)\bigg]\Bigg] + (n-m)\Bigg[\theta B'(sa^*_{I\setminus J}(\theta))a^*_{I\setminus J}(\theta) \\
&- \bigg[a^*_{I\setminus J}(\theta)c\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big) \\
&+ sa^*_{I\setminus J}(\theta)c'\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big) \\
&\cdot \big(ma^*_J(\theta) + (n-m)a^*_{I\setminus J}(\theta)\big)\bigg]\Bigg] + \lambda_1 - \lambda_2 = 0 \\
\lambda_1 s =& 0, \quad \lambda_1 \geq 0, \quad s \geq 0 \\
\lambda_2(1-s) =& 0, \quad \lambda_2 \geq 0, \quad s \leq 1
\end{aligned}$$

<u>Kuhn-Tucker conditions for proportional reduction by J-coalition</u>

$$\max_{\hat{s}\in[0,1]} \; m\left[B(sa^*_J(\theta)) - sa^*_J(\theta)c\Big(msa^*_J(\theta) + (n-m)a^J_{I\setminus J}(\theta)\Big)\right]$$

Maximize the Lagrangian

$$\begin{aligned}
\mathcal{L}(\hat{s}, \lambda_1, \lambda_2) = \; & m\left[B(\hat{s}a^*_J(\theta)) - \hat{s}a^*_J(\theta)c\Big(m\hat{s}a^*_J(\theta) + (n-m)a^J_{I\setminus J}(\theta)\Big)\right] \\
& + \lambda_1\hat{s} + \lambda_2(1-\hat{s})
\end{aligned}$$

and solving the Kuhn-Tucker conditions:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \hat{s}} =& m\Bigg[B'(\hat{s}a^*_J(\theta))a^*_J(\theta) - \bigg[a^*_J(\theta)c\Big(m\hat{s}a^*_J(\theta) + (n-m)a^J_{I\setminus J}(\theta)\Big) \\
&+ \hat{s}a^*_J(\theta)c'\Big(m\hat{s}a^*_J(\theta) + (n-m)a^J_{I\setminus J}(\theta)\Big)\left(ma^*_J(\theta)\right)\bigg]\Bigg] \\
\lambda_1\hat{s} =& 0, \quad \lambda_1 \geq 0, \quad \hat{s} \geq 0 \\
\lambda_2(1-\hat{s}) =& 0, \quad \lambda_2 \geq 0, \quad \hat{s} \leq 1
\end{aligned}$$

Proof of Lemma 1.5.

Restatement of Lemma 1.5. *For any θ, the proportional reduction taken by the grand coalition is never full-reduction, i.e. $s_I^*(\theta) > 0$.*

Proof. Suppose not. Suppose that the coalition of the whole took action of zero. This would mean that the corner solution Kuhn-Tucker condition would have to hold. This would mean that the zero corner slackness multiplier $\lambda_1 \geq 0$, while the one corner slackness multiplier $\lambda_2 = 0$.

The Kuhn-Tucker conditions are:

$$m\Big[B'((0)a_J^*(\theta))a_J^*(\theta) - \Big[a_J^*(\theta)^2 c\Big(m(0)a_J^*(\theta) + (n-m)(0)a_{I\setminus J}^*(\theta)\Big)$$
$$+(0)a_J^*(\theta)^2 c'\Big(m(0)a_J^*(\theta) + (n-m)(0)a_{I\setminus J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\right)\Big]\Big]$$
$$+(n-m)\Big[\theta B'((0)a_{I\setminus J}^*(\theta))a_{I\setminus J}^*(\theta) - \Big[a_{I\setminus J}^*(\theta)^2 c\Big(m(0)a_J^*(\theta) + (n-m)(0)a_{I\setminus J}^*(\theta)\Big)$$
$$+(0)a_{I\setminus J}^*(\theta)^2 c'\Big(m(0)a_J^*(\theta) + (n-m)(0)a_{I\setminus J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\right)\Big]\Big]$$
$$+\lambda_1 = 0$$

These can be simplified as:

$$m\Big[B'(0)a_J^*(\theta) - \Big[a_J^*(\theta)^2 c\,(0) + 0\Big]\Big]$$
$$+ (n-m)\Big[\theta B'(0)a_{I\setminus J}^*(\theta) - \Big[a_{I\setminus J}^*(\theta)^2 c\,(0) + 0\Big]\Big] + \lambda_1 = 0$$

$$\left[ma_J^*(\theta) + (n-m)\theta a_{I\setminus J}^*(\theta)\right] B'(0)$$
$$- \left[ma_J^*(\theta)^2 + (n-m)a_{I\setminus J}^*(\theta)^2\right] c\,(0) + \lambda_1 = 0$$

Since $B'(0) > 0$, $c(0) = 0$, and $\lambda_1 \geq 0$, this equation can only be equal to zero if all the actions are zero. However, we showed that $a_J^*(\theta) > a_{I\setminus J}^*(\theta)$ when $\theta < 1$, so they cannot all be zero. Hence, the grand coalition will never take full-reduction.

$\square$

<u>Incentive for J-coalition to Reduce.</u>

Let the utility of coalition J taking a proportional reduction of s be denoted as

$$u_J^{Pr}(s,\theta) = m\left[B(sa_J^*(\theta)) - sa_J^*(\theta)c\left(msa_J^*(\theta) + (n-m)a_{I\setminus J}^J(\theta)\right)\right].$$

Lemma A.4. *When moving from one, the players in J in the game in Section 1.4.4 have a strict incentive to decrease s from 1 and reduce. Formally, this means that $\left.\frac{\partial u_J^{Pr}(s,\theta)}{\partial s}\right|_{s=1}$, derived below, is negative:*

$$\begin{aligned}\left.\frac{\partial u_J^{Pr}(s,\theta)}{\partial s}\right|_{s=1} =& ma_J^*(\theta)\Big[B'\left(a_J^*(\theta)\right) - c\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^J(\theta)\right) \\ & -mc'\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^J(\theta)\right)\Big]\end{aligned}$$

Proof. This proof is very similar to the proof of Lemma A.3. From the first order conditions, $a_J^*(\theta)$ solves:

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

At $s = 1$, those in J are agreeing to play the no-coalition equilibrium, meaning that the best responses of those not in J are also Nash, giving $a^*_{I\setminus J}(\theta) = a^J_{I\setminus J}(\theta)$. The first order condition and the derivative of u^{Pr}_J with respect to s evaluated at one are then nearly identical, apart from the extra weight on the cost derivative. Comparing the two, then it must be that:

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

Therefore, we have that $\left.\frac{\partial u^{Pr}_J(s,\theta)}{\partial s}\right|_{s=1} < 0$, which means that increasing s will decrease utility – but that decreasing s, thereby increasing reduction, will increase coalition utility. □

Proof of Lemma 1.6.

Restatement of Lemma 1.6. *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 reduction chosen by the coalition of the whole, $s^*_I(\theta)$, is equal to one.*

Proof. Examine the Kuhn-Tucker conditions for s_I.

$$
\begin{aligned}
&ma_J^*(\theta)\Big[B'(sa_J^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\\
&\quad -sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\right)\Big]\\
&\quad +(n-m)a_{I\setminus J}^*(\theta)\Big[\theta B'(sa_{I\setminus J}^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\\
&\quad -sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\right)\Big]\\
&\qquad + \lambda_1 - \lambda_2 = 0
\end{aligned}
$$

According to Lemma 1.5, we know that $s > 0$, so we can ignore one case (and consequently we know that $\lambda_1 = 0$). Let's look at the remaining two cases: interior and corner $s = 1$.

If s were interior, then we would have $\lambda_2 = 0$ as well, and the following would be the K-T condition:

$$
\begin{aligned}
&ma_J^*(\theta)\Big[B'(sa_J^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\\
&\quad -sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\right)\Big]\\
&\quad +(n-m)a_{I\setminus J}^*(\theta)\Big[\theta B'(sa_{I\setminus J}^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\\
&\quad -sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\right)\Big] = 0
\end{aligned}
$$

Basically, this equation is balancing the marginal utility of reduction of the two groups, somehow imagined as $(-MU_J^I) + (-MU_{I\setminus J}^I) = 0$. If both groups were the same, then the optimal solution for the coalition would be the same as for each group. However, since the groups are different, the marginal utilities must take opposite signs to make the equation hold. Therefore, the optimal reduction for the coalition of the whole will have negative marginal utility for one group and positive marginal utility for the other. Moving the terms belonging to $I \setminus J$, the equation becomes:

$$
\begin{aligned}
ma_J^*(\theta)\Big[&B'(sa_J^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big) \\
&-sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\right)\Big] = \\
&-(n-m)a_{I\setminus J}^*(\theta)\Big[\theta B'(sa_{I\setminus J}^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big) \\
&-sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\right)\Big]
\end{aligned}
$$

Since $a_{I\setminus J}^*(\theta)$ is continuous and increasing in $\theta \Rightarrow \lim_{\theta\to 0} a_{I\setminus J}^*(\theta) = 0.$[1] Consider an action close to zero, particularly $a_{I\setminus J}^*(\theta) = \frac{\varepsilon}{n-m} > 0$, but very small. Then it looks like:

$$
\begin{aligned}
ma_J^*(\theta)&\Big[B'(sa_J^*(\theta)) - c(msa_J^*(\theta) + s\varepsilon) - sc'(msa_J^*(\theta) + s\varepsilon)\left(ma_J^*(\theta) + \varepsilon\right)\Big] = \\
&-\varepsilon\left[\theta B'\left(s\frac{\varepsilon}{n-m}\right) - c(msa_J^*(\theta) + s\varepsilon) - sc'(msa_J^*(\theta) + s\varepsilon)\left(ma_J^*(\theta) + \varepsilon\right)\right]
\end{aligned}
$$

[1] If the parameter set included $\theta = 0$, then at the action $a_{I\setminus J}^*(1) = 0$ the equation would look like:

$$
ma_J^*(\theta)\Big[B'(sa_J^*(\theta)) - c(msa_J^*(\theta)) - sc'(msa_J^*(\theta))ma_J^*(\theta)\Big] = 0
$$

This could actually have a positive solution for s, so literal zero producers are permitted to hang on. However, the parameter set does not include zero.

Because of Nash equilibrium, we know that:

$$B'(a_J^*(\theta)) - c(ma_J^*(\theta) + \varepsilon) - mc'(ma_J^*(\theta) + \varepsilon) = 0$$

If s is interior, then $sa_J^*(\theta) < a_J^*(\theta)$. This means that $B'(sa_J^*(\theta)) \geq B'(a_J^*(\theta))$, and $c(msa_J^*(\theta) + s\varepsilon) < c(ma_J^*(\theta) + \varepsilon)$. The remaining term appears ambiguous at first:

$$\begin{aligned}
mc'(ma_J^*(\theta) + \varepsilon) &\overset{?}{>} s\,(ma_J^*(\theta) + \varepsilon)\,c'(msa_J^*(\theta) + s\varepsilon)\\
mc'(ma_J^*(\theta) + \varepsilon) &\overset{?}{>} msa_J^*(\theta)c'(msa_J^*(\theta) + s\varepsilon) + s\varepsilon c'(msa_J^*(\theta) + s\varepsilon)\\
m\left[c'(ma_J^*(\theta) + \varepsilon) - sa_J^*(\theta)c'(msa_J^*(\theta) + s\varepsilon)\right] &\overset{?}{>} s\varepsilon c'(msa_J^*(\theta) + s\varepsilon)
\end{aligned}$$

To make LHS > 0, choose

$$\varepsilon < \frac{m\left[c'(ma_J^*(\theta) + \varepsilon) - sa_J^*(\theta)c'(msa_J^*(\theta) + s\varepsilon)\right]}{sc'(msa_J^*(\theta) + s\varepsilon)}.$$

From the Nash equilibrium, we also know that:

$$\theta B'\left(\frac{\varepsilon}{n-m}\right) - c(ma_J^*(\theta) + \varepsilon) - (n-m)c'(ma_J^*(\theta) + \varepsilon) = 0$$

If the marginal benefit of action for those not in J, i.e. $\theta B'\left(s\frac{\varepsilon}{n-m}\right)$, is large enough (larger than the cost and marginal cost), then RHS is negative. By similar logic as above, this should hold. Hence, the multiplier λ_2 needs to be included and greater than zero in order to balance out the FOC, meaning that $s = 1$ is chosen when θ is small enough.

$\square$

Appendix B

Alternative Proofs for Chapter 1

B.1 Lump-Sum Reduction under Stronger Assumptions

If we consider a few stronger assumptions on the class of negative externality games, then we can still demonstrate openness of unique Nash equilibrium and, hence, the exclusion result. Consider $\mathcal{U}'$ 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, $(\forall i \in I)(\forall a \in A)\left[\left|\frac{\partial^2 u_i(a)}{\partial a_i^2}\right| > \left|\sum_{j \neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j}\right|\right]$, and

f. unique interior Nash equilibrium, the Kuhn-Tucker conditions for equilibrium of u have a unique solution on the interior of A.

Observe here, that the stronger assumptions are unique *interior* Nash equilibrium, as well as the added assumption of strong dominant effect. The proofs in Appendix A did not require these two stronger assumptions. It is a useful exercise to run through the details under these stronger assumptions.

Use the same utility function as before:

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

The function's fulfillment of twice continuous differentiability, negative externalities, submodularity, and strict own concavity are shown in Appendix A. Furthermore, while showing unique Nash, the only equilibrium was shown to be interior as well. The only remaining property to check is 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.

<u>Proof of Openness of $\mathcal{U}'$.</u>

Proof. Again, 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 openness of the conditions of negative externalities, submodularity, and concavity under this norm was shown in Appendix A. The remaining two conditions are conditionally open, i.e. open when the previous three conditions hold. To show conditional openness for the two new conditions, 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.

- *Strong dominant effect:* The previous three properties were unconditionally open. For the uses of Theorem 1.1, it is sufficient for strong dominant effect to be open conditional upon the previous three properties.

  Lemma B.1. *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 $\sum_{j\neq i} \frac{\partial^2 v_i(a)}{\partial a_i \partial a_j}$:

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

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

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

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

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

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

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:

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

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:* For the purposes of Theorem 1.1, it is sufficient for the property of unique interior Nash equilibrium to be open conditional upon the previous four properties.

Lemma B.2. *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 [22]: *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:

- $X = \mathfrak{U}'$
- $Y = [0,1]^n$
- $Z = \mathbb{R}^n$
- $\Omega \subseteq (0,1)^n$
- $F = \begin{pmatrix} \frac{\partial v_1(a)}{\partial a_1} \\ \frac{\partial v_2(a)}{\partial a_2} \\ \vdots \\ \frac{\partial v_n(a)}{\partial a_n} \end{pmatrix}$
- (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:

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

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 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, either unconditionally or conditionally upon the remaining conditions. 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.

□

B.2 Proportional Reduction under Stronger Assumptions

To establish the exclusion result under proportional commitment power, I add two conditions on the utility functions in $\mathcal{U}'$, creating a subset $\mathcal{U}''$. In addition to the previous six conditions for $\mathcal{U}'$, the utility functions belong to the class $\mathcal{U}''$ must satisfy the following conditions:

g. infinite marginal benefit at zero, $\lim_{a_i \to 0} \frac{\partial u_i(a)}{\partial a_i} = \infty$, and

h. finite damages, $\lim_{a \to \mathbf{1}} \left| \frac{\partial u_i(a)}{\partial a} \right| < \infty$.

These additional conditions strengthen the reasoning of disparate marginal utilities and give sufficiency for the exclusion result under proportional reduction. The conditions are possibly reasonable in the context of a negative externality situation. The first condition stipulates that the action which causes the negative externality is necessary in some manner. It may be unthinkable for a country to completely terminate an industry that releases pollutants or a fishery that provides much of a region's food. Unique interior Nash equilibrium was already one of the assumptions on $\mathcal{U}$; now the reason for non-zero action is the infinite marginal benefit at zero. Meanwhile, the second new condition simply means that costs from action are not infinite at the maximum possible stock level, where every agent plays one. This emphasizes the power of the negative externality as the reason the Nash equilibrium is interior from the other side.

Alternative Proof of Existence of Exclusion under Proportional Reductions.

Theorem B.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 $a^{Pr}(u, J)$ is conceivable, formally denoted as:*

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

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

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

Proof. Again, we use the same approach as for the other proofs, this time showing that the two additional conditions are fulfilled and open.

The new chosen utility function is:

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

Again, the parameter θ is as described in Section 1.4.2, where the group J consists of agents who have $\theta_i = 0$, while the remaining agents not in J have $\theta_i = \theta \in \Theta = (0, 1)$.

Exclusion Result

The optimality of exclusion comes from Lemma B.4, which is itself proven later in this section.

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

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

  - Second Derivatives

$$\begin{aligned} \tfrac{d^2u_i(a)}{da^2} &= \tfrac{\partial^2 u_i(a)}{\partial a_i^2} + 2\textstyle\sum_{j\neq i} \tfrac{\partial^2 u_i(a)}{\partial a_i \partial a_j} + \textstyle\sum_{j\neq i}\left\{\tfrac{\partial^2 u_i(a)}{\partial a_j^2} + \textstyle\sum_{k\neq i \,\mathrm{or}\, j} \tfrac{\partial u_i(a)}{\partial a_j \partial a_k}\right\} \\ \tfrac{\partial^2 u_i(a)}{\partial a_i^2} &= \tfrac{-(1-\theta_i)}{a_i^2} - \left[2\left(n + \textstyle\sum_{j=1}^{n} a_j\right) + 4a_i\right] \end{aligned}$$

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

$$\forall\, j \neq i \;\; \tfrac{\partial u_i(a)}{\partial a_j^2} = 0$$

$$\forall\, k \neq i \,\mathrm{or}\, j \;\; \tfrac{\partial u_i(a)}{\partial a_j \partial a_k} = 0$$

- *Negative externalities:* The first derivative of i's utility function with respect to any j's action is the same as in the proof of Theorem 1.1:

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

  Once again, 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:* The cross-partial of i's utility function with respect to his own action and another agent j's action is the same as in the proof of Theorem 1.1:

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

  Again, 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:* The second derivative of i's utility function with respect to his own action is:

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

  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 $j \neq i$.

$$\left|\frac{\partial^2 u_i(a)}{\partial a_i^2}\right| \overset{?}{>} \left|\sum_{j\neq i} \frac{\partial^2 u_i(a)}{\partial a_i \partial a_j}\right|$$

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

$$\frac{(1-\theta_i)}{a_i^2} + \left[2\left(n + \sum_{j=1}^n a_j\right) + 4a_i\right] \overset{?}{>} 2(n-1)a_i$$

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

As with Theorem 1.1, because $n > n-1$ and $a_i \leq 1 \Rightarrow 2n > 2(n-1)a_i$, and $\frac{(1-\theta_i)}{a_i^2} + 2\sum_{j=1}^n a_j + 4a_i \geq 0 \Rightarrow$

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

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

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

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

The Lagrangian is:

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

The Kuhn-Tucker conditions are:

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

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:

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

This cubic equation can be easily solved as a function of $\sum_{j\neq i} a_j^*$ when n and θ_i are known, but without these two parameters, the closed form is difficult. However, the roots can be characterized using the polynomial discriminant and Descartes' Rule of Signs.

- **Polynomial discriminant:** A trinomial represented by $ax^3 + bx^2 + cx + d = 0$ has the discriminant:

$$\Delta = b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd$$

Therefore, for the polynomial in question,

$$3a_i^3 + 2a_i^2\left(n + \sum_{j\neq i} a_j^*\right) - (1-\theta_i) = 0,$$

the discriminant is:

$$
\begin{aligned}
\Delta = &\left(2\left(n+\textstyle\sum_{j\neq i} a_j^*\right)\right)^2 (0)^2 - 4(3)(0)^3 \\
&-4\left(2\left(n+\textstyle\sum_{j\neq i} a_j^*\right)\right)^3 (-(1-\theta_i)) - 27(3)^2\left(-(1-\theta_i)\right)^2 \\
&+18(3)\left(2\left(n+\textstyle\sum_{j\neq i} a_j^*\right)\right)(0)\left(-(1-\theta_i)\right) \\
= &-4\left(2\left(n+\textstyle\sum_{j\neq i} a_j^*\right)\right)^3 (-(1-\theta_i)) - 27(3)^2\left(-(1-\theta_i)\right)^2 \\
= &\,4(8)\left(n+\textstyle\sum_{j\neq i} a_j^*\right)^3 (1-\theta_i) - 27(9)(1-\theta_i)^2 \\
= &\,32\left(n+\textstyle\sum_{j\neq i} a_j^*\right)^3 (1-\theta_i) - 243(1-\theta_i)^2
\end{aligned}
$$

A trinomial has possibly three roots. Knowing the sign of the discriminant can aid in determining the nature of those roots. This discriminant will be positive when:

$$
\begin{aligned}
32\left(n+\textstyle\sum_{j\neq i} a_j^*\right)^3 (1-\theta_i) - 243(1-\theta_i)^2 > 0 \\
32\left(n+\textstyle\sum_{j\neq i} a_j^*\right)^3 (1-\theta_i) > 243(1-\theta_i)^2 \\
32\left(n+\textstyle\sum_{j\neq i} a_j^*\right)^3 > 243(1-\theta_i)
\end{aligned}
$$

The term $(1-\theta_i)$ is smaller than one, so the following is sufficient:

$$
\begin{aligned}
32\left(n+\textstyle\sum_{j\neq i} a_j^*\right)^3 &> 243 \\
\left(n+\textstyle\sum_{j\neq i} a_j^*\right)^3 &> \frac{243}{32} \\
n^3 + 3n^2\textstyle\sum_{j\neq i} a_j^* + 3n\left(\textstyle\sum_{j\neq i} a_j^*\right)^2 + \left(\textstyle\sum_{j\neq i} a_j^*\right)^3 &> \frac{243}{32}
\end{aligned}
$$

If a_{-i}^* is the zero vector, then $n^3 > \frac{243}{32}$ is sufficient for the discriminant to be positive. Two facts guarantee that this is true: first, $n > 2$, which means that $n^3 > 8$, while $\frac{243}{32} < 8$. If a^*_{-i} is greater than the zero vector, then the condition on n is even smaller. Hence, the discriminant for this polynomial is always positive, meaning that there are three distinct real roots.

– **Descartes' Rule of Signs:** Looking at the polynomial discriminant gives three candidate solutions. Descartes' Rule of Signs can characterize how many of of these are positive, narrowing down the field of maximizers.

  The polynomial derived from the interior Kuhn-Tucker condition gives the function:

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

According to Descartes' Rule of Signs, the number of sign switches between non-zero coefficients gives the maximum number of positive roots. Here, we have one sign switch, since the first and second terms are positive but the third term is negative. Therefore, there is at most one positive root out of the three real roots. We know this root is greater than zero; we will check whether it is less than one when checking the possibility of a corner solution at one. This root will be the interior solution $a_i^*(\theta_i, a^*_{-i})$.

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

The first derivative of the Lagrangian becomes:

$$\frac{(1-\theta_i)}{a_i} + \lambda_{1i} = 0$$

Since a_i is in the denominator it cannot be zero. Therefore, this case will never occur.

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

The first derivative of the Lagrangian becomes:

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

However, since $1-\theta_i \leq 1$ and $2\left(1 + \sum_{j\neq i} a_j^*\right) >> 1$, this case can never occur. Because of this, we know that the earlier interior case is less than one.

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:

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

These can be rewritten as:

$$2a_i^{*2}\left(n + \sum_{j=1}^n a_j^*\right) + a_i^{*3} = (1-\theta_i)$$
$$2a_i^{**2}\left(n + \sum_{j=1}^n a_j^{**}\right) + a_i^{**3} = (1-\theta_i)$$

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

$$2\left(n + \sum_{i=1}^{n} a_i^*\right) \sum_{i=1}^{n} a_i^{*2} + \sum_{i=1}^{n} a_i^{*3} = \sum_{i=1}^{n}(1-\theta_i)$$

$$2\left(n + \sum_{i=1}^{n} a_i^{**}\right) \sum_{i=1}^{n} a_i^{**2} + \sum_{i=1}^{n} a_i^{**3} = \sum_{i=1}^{n}(1-\theta_i)$$

Subtract the bottom condition from the top one:

$$\left[2\left(n + \sum_{i=1}^{n} a_i^*\right) \sum_{i=1}^{n} a_i^{*2} + \sum_{i=1}^{n} a_i^{*3}\right]$$
$$- \left[2\left(n + \sum_{i=1}^{n} a_i^{**}\right) \sum_{i=1}^{n} a_i^{**2} + \sum_{i=1}^{n} a_i^{**3}\right] = 0$$

$$2\left[\left(n + \sum_{i=1}^{n} a_i^*\right) \sum_{i=1}^{n} a_i^{*2} - \left(n + \sum_{i=1}^{n} a_i^{**}\right) \sum_{i=1}^{n} a_i^{**2}\right]$$
$$+ \left[\sum_{i=1}^{n} a_i^{*3} - \sum_{i=1}^{n} a_i^{**3}\right] = 0$$

$$2n\left[\sum_{i=1}^{n} a_i^{*2} - \sum_{i=1}^{n} a_i^{**2}\right] + 2\left[\left(\sum_{i=1}^{n} a_i^*\right) \sum_{i=1}^{n} a_i^{*2} - \left(\sum_{i=1}^{n} a_i^{**}\right) \sum_{i=1}^{n} a_i^{**2}\right]$$
$$+ \left[\sum_{i=1}^{n} a_i^{*3} - \sum_{i=1}^{n} a_i^{**3}\right] = 0$$

All three of these terms will have the same sign, since in between zero and one

$$a_i^* > a_i^{*2} > a_i^{*3}.$$

Adding these up, we know that:

$$\sum_{i=1}^{n} a_i^* > \sum_{i=1}^{n} a_i^{*2} > \sum_{i=1}^{n} a_i^{*3}.$$

Furthermore, since on $(0,1)$, $f(x) = x^2$ is strictly increasing, we have that:

$$a_i^* > a_i^{**} \Rightarrow a_i^{*2} > a_i^{**2}.$$

Adding these up, we know that:

$$\sum_{i=1}^{n} a_i^* > \sum_{i=1}^{n} a_i^{**} \Rightarrow \sum_{i=1}^{n} a_i^{*2} > \sum_{i=1}^{n} a_i^{**2}.$$

Suppose that $\sum_{i=1}^{n} a_i^* > \sum_{i=1}^{n} a_i^{**}$. This means that the first and third terms are positive. The middle term is $x \cdot y - x' \cdot y'$ where we know that $x > x'$ and $y > y'$, so we know that it is positive as well. Adding three positive terms cannot give zero.

Now suppose that $\sum_{i=1}^{n} a_i^* < \sum_{i=1}^{n} a_i^{**}$. This means that all three terms are negative. Adding three negative terms cannot give zero either.

Therefore, this equation can only be solved if each of the three terms is zero, and we have that $\sum_{i=1}^{n} a_i^* = \sum_{i=1}^{n} a_i^{**}$, which also gives that $\sum_{i=1}^{n} a_i^{*2} = \sum_{i=1}^{n} a_i^{**2}$ and $\sum_{i=1}^{n} a_i^{*3} = \sum_{i=1}^{n} a_i^{**3}$. 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.

$$2a_i^{*2}\left(n + \sum_{j=1}^{n} a_j^*\right) + a_i^{*3} = (1 - \theta_i)$$
$$2a_i^{**2}\left(n + \sum_{j=1}^{n} a_j^{**}\right) + a_i^{**3} = (1 - \theta_i)$$

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:

$$2(a_i^{*2} - a_i^{**2})\left(n + \sum_{j=1}^{n} a_j^*\right) + (a_i^{*3} - a_i^{**3}) = 0$$

Substitute the factorizations for the difference of squares and difference of cubes:

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

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:

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

This leads 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.

- *Infinite marginal benefit at zero:* The first derivative of i's utility with respect to a_i is:

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

The limit as a_i approaches zero is:

$$\begin{aligned}
\lim_{a_i \to 0} \frac{\partial u_i(a)}{\partial a_i} &= \lim_{a_i \to 0}\left\{\tfrac{(1-\theta_i)}{a_i} - \left[2a_i\left(n + \sum_{j=1}^n a_j\right) + a_i^2\right]\right\} \\
&= \lim_{a_i \to 0}\tfrac{(1-\theta_i)}{a_i} - \lim_{a_i \to 0}\left[2a_i\left(n + \sum_{j=1}^n a_j\right) + a_i^2\right] \\
&= \infty + 0 \\
&= \infty
\end{aligned}$$

Therefore, the utility function has infinite marginal benefit at zero.

- *Finite damages:* The total derivative of i's utility with respect to the vector a is:

$$\begin{aligned}\tfrac{du_i(a)}{da} &= \tfrac{\partial u_i(a)}{\partial a_i} + \textstyle\sum_{j\neq i} \tfrac{\partial u_i(a)}{\partial a_j} \\ &= \tfrac{(1-\theta_i)}{a_i} - \left[2a_i\left(n + \textstyle\sum_{j=1}^{n} a_j\right) + a_i^2\right] - \sum_{j\neq i} a_i^2\end{aligned}$$

The limit of the absolute value as a approaches $\mathbf{1}$ (each element of the vector a approaches 1) is:

$$\begin{aligned}\lim_{a\to\mathbf{1}}\left|\frac{\partial u_i(a)}{\partial a}\right| &= \lim_{a\to\mathbf{1}}\left|\tfrac{(1-\theta_i)}{a_i} - \left[2a_i\left(n + \textstyle\sum_{j=1}^{n} a_j\right) + a_i^2\right] - \sum_{j\neq i} a_i^2\right| \\ &= \lim_{a\to\mathbf{1}}\left|\tfrac{(1-\theta_i)}{a_i}\right| + \lim_{a\to\mathbf{1}}\left|\left[2a_i\left(n + \textstyle\sum_{j=1}^{n} a_j\right) + a_i^2\right]\right| \\ &+ \lim_{a\to\mathbf{1}}\left|\textstyle\sum_{j\neq i} a_i^2\right| \\ &= (1-\theta_i) + [4n+1] + (n-1)^2\end{aligned}$$

This limit is a finite number, therefore finite damages holds.

Openness of Conditions

In Appendix B.1, I have already shown that the properties of negative externality, submodularity, and concavity are unconditionally open, that the property of strong dominant effect is open conditional upon the previous three, and that the property of unique interior Nash equilibrium is open conditional upon the presence of all four preceding properties. Now I will show openness of the additional properties of $\mathcal{U}''$.

- *Infinite marginal benefit at zero:* Define $\varepsilon_e = \frac{1}{2}$. If $||u - v|| < \varepsilon_e$, then for any i, for any a, $\left|\frac{\partial u_i(a)}{\partial a_i} - \frac{\partial v_i(a)}{\partial a_i}\right| < \varepsilon_e$. At each a, even when a_i is very small and getting, the distance between these derivatives is at most ε_e. This means that $\lim_{a_i \to 0} \left| \frac{\partial u_i(a)}{\partial a_i} - \frac{\partial v_i(a)}{\partial a_i} \right| < \varepsilon_e$. Since $u \in \mathcal{U}$, we know that $\lim_{a_i \to 0} \frac{\partial u_i(a)}{\partial a_i} = \infty$, so it must be the case that $\lim_{a_i \to 0} \frac{\partial v_i(a)}{\partial a_i} = \infty$ as well. Hence, v has infinite marginal benefit at zero.

- *Finite damages:* The problem $\min_i \left| \frac{\partial u_i(a)}{\partial a} \right|_{a=\mathbf{1}}$ has a solution $\{i^*\}$. Since $u \in \mathcal{U}'$, u fulfills finite damages, so we know that $\left| \frac{\partial u_i(a)}{\partial a} \right| < \infty$. This value is either the smallest positive total derivative or the largest negative total derivative. Define

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

  If $||u - v|| < \varepsilon_f$, then $\forall i \; \frac{\partial v_i(\mathbf{1})}{\partial a} \in \left(\frac{\partial u_i(\mathbf{1})}{\partial a} - \varepsilon_f, \frac{\partial u_i(\mathbf{1})}{\partial a} + \varepsilon_f \right)$. These are two finite values, and since $\frac{\partial v_i(\mathbf{1})}{\partial a}$ is between them, it too is finite. Hence, $\lim_{a \to \mathbf{1}} \left| \frac{\partial v_i(a)}{\partial a} \right| < \infty$, and v has finite damages.

Each of the original and additional conditions has been shown to be open. Take

$$\varepsilon^* \equiv \min\{\varepsilon_a, \varepsilon_b, \varepsilon_c, \varepsilon_d, \varepsilon_e, \varepsilon_f\}.$$

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.

$\square$

Proportional Reduction Lemmas in $\mathcal{U}''$.

Lemma B.3. *For any θ, the proportional reduction taken by the grand coalition is never full-reduction, i.e. $s_I^*(\theta) > 0$.*

Proof. Suppose not. Suppose that the coalition of the whole took action of zero. This would mean that the corner solution Kuhn-Tucker condition would have to hold. This would mean that the zero corner slackness multiplier $\lambda_1 \geq 0$, while the one corner slackness multiplier $\lambda_2 = 0$.

The Kuhn-Tucker conditions are:

$$m\Big[B'((0)a_J^*(\theta))a_J^*(\theta) - \Big[a_J^*(\theta)^2 c\Big(m(0)a_J^*(\theta) + (n-m)(0)a_{I\setminus J}^*(\theta)\Big)$$
$$+ (0)a_J^*(\theta)^2 c'\Big(m(0)a_J^*(\theta) + (n-m)(0)a_{I\setminus J}^*(\theta)\Big)$$
$$\cdot \big(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\big)\Big]\Big] + (n-m)\Big[(1-\theta)B'((0)a_{I\setminus J}^*(\theta))a_{I\setminus J}^*(\theta)$$
$$- \Big[a_{I\setminus J}^*(\theta)^2 c\Big(m(0)a_J^*(\theta) + (n-m)(0)a_{I\setminus J}^*(\theta)\Big)$$
$$+ (0)a_{I\setminus J}^*(\theta)^2 c'\Big(m(0)a_J^*(\theta) + (n-m)(0)a_{I\setminus J}^*(\theta)\Big)$$
$$\cdot \big(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\big)\Big]\Big] + \lambda_1 = 0$$

However, since $\lim_{a_i \to 0} B'(a_i) = \infty$, this equation is adding two positive infinite terms, which cannot sum to zero. Therefore, because of the condition of infinite marginal benefit at zero, the grand coalition will never take full-reduction.

$\square$

Lemma B.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 reduction chosen by the coalition of the whole, $s_I^*(\theta)$, is equal to one.*

Proof. Examine the Kuhn-Tucker conditions for s_I.

$$\begin{aligned}
& ma^*_J(\theta)\Big[B'(sa^*_J(\theta)) - c\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big) \\
& \quad -sc'\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big)\left(ma^*_J(\theta) + (n-m)a^*_{I\setminus J}(\theta)\right)\Big] \\
& \quad +(n-m)a^*_{I\setminus J}(\theta)\Big[(1-\theta)B'(sa^*_{I\setminus J}(\theta)) - c\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big) \\
& \quad -sc'\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big)\left(ma^*_J(\theta) + (n-m)a^*_{I\setminus J}(\theta)\right)\Big] \\
& \quad +\lambda_1 - \lambda_2 = 0
\end{aligned}$$

According to Lemma B.3, we know that $s > 0$, so we can ignore one case (and consequently we know that $\lambda_1 = 0$). Let's look at the remaining two cases: interior and corner $s = 1$.

If s were interior, then we would have $\lambda_2 = 0$ as well, and the following would be the K-T condition:

$$\begin{aligned}
& ma^*_J(\theta)\Big[B'(sa^*_J(\theta)) - c\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big) \\
& \quad -sc'\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big)\left(ma^*_J(\theta) + (n-m)a^*_{I\setminus J}(\theta)\right)\Big] \\
& \quad +(n-m)a^*_{I\setminus J}(\theta)\Big[(1-\theta)B'(sa^*_{I\setminus J}(\theta)) - c\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big) \\
& \quad -sc'\Big(msa^*_J(\theta) + (n-m)sa^*_{I\setminus J}(\theta)\Big)\left(ma^*_J(\theta) + (n-m)a^*_{I\setminus J}(\theta)\right)\Big] = 0
\end{aligned}$$

Basically, this equation is balancing the marginal utility of reduction of the two groups, somehow imagined as $(-MU^I_J) + (-MU^I_{I\setminus J}) = 0$. If both groups were the same, then the optimal solution for the coalition would be the same as for each group. However, since the groups are different, the marginal utilities must take opposite signs to make the equation hold. Therefore, the optimal reduction for the coalition of the whole will have negative marginal utility for one group and positive marginal utility for the other. Moving the terms belonging to $I \setminus J$, the equation becomes:

$$
\begin{aligned}
ma_J^*(\theta)\Big[B'(sa_J^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big) \\
-sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\right)\Big] = \\
-(n-m)a_{I\setminus J}^*(\theta)\Big[(1-\theta)B'(sa_{I\setminus J}^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big) \\
-sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\setminus J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\setminus J}^*(\theta)\right)\Big]
\end{aligned}
$$

Since $a_{I\setminus J}^*(\theta)$ is continuous and decreasing in $\theta \Rightarrow \lim_{\theta\to 1} a_{I\setminus J}^*(\theta) = 0$. Since $u \in \mathcal{U}'$, it must be that $\lim_{a_i\to 0} \frac{\partial u_i(a)}{\partial a_i} = \infty$, so for this utility function that means that $\lim_{\theta\to 1, a_i\to 0}(1-\theta)B'(a_i) = \infty$, i.e. $B'(\cdot)$ goes to infinity faster than θ goes to one. Therefore, as θ approaches one, RHS is going to negative infinity quickly. In fact, with a small, tiny s close to zero multiplying the action, then RHS would go to infinity *even faster*. In order for the equation to hold, LHS must to go to negative infinity as well. The benefit function, $B'(\cdot)$, is strictly increasing throughout the action space, so it cannot make RHS go to negative infinity. The cost function is subtracted and has the potential to make RHS negative. However, by the assumption made that the cost function does not asymptote to infinity within the action space of $[0,1]$, this means that neither $c(\cdot)$ nor $c'(\cdot)$ goes to infinity. Therefore, RHS cannot go to negative infinity, so the two cannot be equal and an interior s is not possible when θ is really, really tiny.

Instead, look at the K-T condition for the corner of $s = 1$, where $\lambda_2 \geq 0$:

$$\begin{aligned}
&ma_J^*(\theta)\Big[B'(sa_J^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\backslash J}^*(\theta)\Big) \\
&\quad -sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\backslash J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\backslash J}^*(\theta)\right)\Big] \\
&\quad +(n-m)a_{I\backslash J}^*(\theta)\Big[(1-\theta)B'(sa_{I\backslash J}^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\backslash J}^*(\theta)\Big) \\
&\quad -sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\backslash J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\backslash J}^*(\theta)\right)\Big] - \lambda_2 = 0
\end{aligned}$$

This can be rewritten as:

$$\begin{aligned}
&ma_J^*(\theta)\Big[B'(sa_J^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\backslash J}^*(\theta)\Big) \\
&\quad -sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\backslash J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\backslash J}^*(\theta)\right)\Big] - \lambda_2 = \\
&\quad -(n-m)a_{I\backslash J}^*(\theta)\Big[(1-\theta)B'(sa_{I\backslash J}^*(\theta)) - c\Big(msa_J^*(\theta) + (n-m)sa_{I\backslash J}^*(\theta)\Big) \\
&\quad -sc'\Big(msa_J^*(\theta) + (n-m)sa_{I\backslash J}^*(\theta)\Big)\left(ma_J^*(\theta) + (n-m)a_{I\backslash J}^*(\theta)\right)\Big]
\end{aligned}$$

Since $\lambda_2 \geq 0$, subtracting it helps balance things, and it is permissible that $\lambda_2 \to \infty$ as $a_{I\backslash J}^* \to 0$. This means that $s_I^*(\theta) = 1$ will occur when θ is large enough.

□

Appendix C

Proofs and Derivations for Chapter 2

C.1 Alternative Parametrization with Centered Taylor Expansions

In this paper, simple linear parameters are used. However, an alternative similar to Taylor expansions was suggested. Here are the respective set ups for the three different effects.

1. Opponent effect:

   (a) Non-coordination:

   $$w_i^N(\theta_J) = u_i(a_i, a_j) - \theta_J(a_j - a_j^N(0))$$
   $$w_j^N(\theta_J) = u_j(a_j, a_i) - \theta_J(a_i - a_i^N(0))$$

   (b) Coordination:

   $$w^P(\theta_J) = u_i(a_i, a_j) + u_j(a_j, a_i) - \theta_J\left[(a_i - a_i^P(0)) + (a_j - a_j^P(0))\right]$$

2. Own effect:

   (a) Non-coordination:

   $$w_i^N(\theta_I) = u_i(a_i, a_j) + \theta_I(a_i - a_i^N(0))$$
   $$w_j^N(\theta_I) = u_j(a_j, a_i) + \theta_I(a_j - a_j^N(0))$$

(b) Coordination:

$$w^P(\theta_I) = u_i(a_i, a_j) + u_j(a_j, a_i) + \theta_I \left[(a_i - a_i^P(0)) + (a_j - a_j^P(0))\right]$$

3. Submodular effect:

(a) Non-coordination:

$$w_i^N(\theta_{IJ}) = u_i(a_i, a_j) - \theta_{IJ} \left(a_i - a_i^N(0)\right) \left(a_j - a_j^N(0)\right)$$
$$w_j^N(\theta_{IJ}) = u_j(a_j, a_i) - \theta_{IJ} \left(a_i - a_i^N(0)\right) \left(a_j - a_j^N(0)\right)$$

(b) Coordination:

$$w_i^P(\theta_{IJ}) = u_i(a_i, a_j) + u_j(a_j, a_i) - 2\theta_{IJ} \left(a_i - a_i^P(0)\right) \left(a_j - a_j^P(0)\right)$$

C.2 Expanded Results from Section 2.3.1

Proof of Theorem 2.1

Restatement of Theorem 2.1 from Section 2.3.1. *For a symmetric game Γ, an increase in the parameter multiplying the added linearizations has the following effect for each:*

1. *Increasing the opponent effect increases the distance in the actions under non-coordination and coordination, that is, for all i:*

$$\frac{d}{d\theta_J}[a_i^N(\theta_J) - a_i^P(\theta_J)] > 0;$$

2. *Increasing the own effect has ambiguous results on the distance in actions under non-coordination and coordination; and*

3. *Increasing the submodular effect also has ambiguous results on the distance in actions under non-coordination and coordination.*

Proof. For each type of effect, this proof examines the first order conditions to determine the directions of change in the action gaps. Each effect has separate analysis.

1. **Opponent Effect**

   The opponent effect is set-up in the paper in Equations (2.4) and (2.5). First, I examine the Nash first order conditions, and then I examine the social planner's first order conditions.

   (a) Non-coordination: The maximization problem for agent i, given the Nash equilibrium action of agent j, is:

   $$\max_{a_i \in A_i} \; u_i(a_i, a_j^N(\theta_J)) - \theta_J a_j^N(\theta_J)$$

   $a^N(0)$ solves the following:

   $$\frac{\partial w_i}{\partial a_i} = \frac{\partial u_i(\cdot, a_j^N(0))}{\partial a_i} \equiv 0$$
   $$\frac{\partial w_j}{\partial a_j} = \frac{\partial u_j(\cdot, a_i^N(0))}{\partial a_j} \equiv 0$$

   $a^N(\theta_J)$ solves the following:

$$\frac{\partial w_i}{\partial a_i} = \frac{\partial u_i(\cdot, a_j^N(\theta_J))}{\partial a_i} \equiv 0$$
$$\frac{\partial w_j}{\partial a_j} = \frac{\partial u_j(\cdot, a_i^N(\theta_J))}{\partial a_j} \equiv 0$$

Observe that $a^N(\theta_J) = a^N(\theta'_J) = a^N(0)$ for all θ_J and θ'_J in θ_J. Hence,

$$\frac{\partial a_i^N(\cdot)}{\partial \theta_J} = 0.$$

(b) Coordination: The maximization problem for the social planner is:

$$\max_{(a_i,a_j)\in A_i\times A_j} u_i(a_i, a_j) + u_j(a_j, a_i) - \theta_J\,(a_i + a_j)$$

For $\theta_J = 0$, $a^P(0)$ solves the following:

$$\frac{\partial(w_i + w_j)}{\partial a_i} = \frac{\partial u_i(\cdot, a_j^P(\theta_J))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_J), \cdot)}{\partial a_i} \equiv 0$$
$$\frac{\partial(w_i + w_j)}{\partial a_j} = \frac{\partial u_j(\cdot, a_i^P(\theta_J))}{\partial a_j} + \frac{\partial u_i(a_i^P(\theta_J), \cdot)}{\partial a_j} \equiv 0$$

For $\theta_J > 0$, $a^P(\theta_J)$ solves the following:

$$\frac{\partial(w_i + w_j)}{\partial a_i} = \frac{\partial u_i(\cdot, a_j^P(\theta_J))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_J), \cdot)}{\partial a_i} - \theta_J \equiv 0$$
$$\frac{\partial(w_i + w_j)}{\partial a_j} = \frac{\partial u_j(\cdot, a_i^P(\theta_J))}{\partial a_j} + \frac{\partial u_i(a_i^P(\theta_J), \cdot)}{\partial a_j} - \theta_J \equiv 0$$

Since the game is symmetric, if the Social Planner changes any agent's action, he will change the other's action in the same manner (i.e same direction and likely magnitude). The next lemma looks at the comparative statics of the whole vector.

Lemma C.1. *$a^P(\theta_J)$ is decreasing in θ_J.*

Proof. Suppose not. Suppose that for $\theta'_J > \theta_J$, $a^P(\theta'_J) \nless a^P(\theta_J)$.

i. Case i. $a^P(\theta'_J) > a^P(\theta_J)$

Look at the FOC for $\frac{\partial(w_i+w_j)}{\partial a_i}$ (the FOC for $\frac{\partial(w_i+w_j)}{\partial a_j}$ are symmetric):

$$\frac{\partial u_i(a_i^P(\theta_J), a_j^P(\theta_J))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_J), a_i^P(\theta_J))}{\partial a_i} = \theta_J$$
$$\frac{\partial u_i(a_i^P(\theta'_J), a_j^P(\theta'_J))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta'_J), a_i^P(\theta'_J))}{\partial a_i} = \theta'_J$$

Subtract the first from the second:

$$\begin{aligned}&\left[\frac{\partial u_i(a_i^P(\theta'_J), a_j^P(\theta'_J))}{\partial a_i} - \frac{\partial u_i(a_i^P(\theta_J), a_j^P(\theta_J))}{\partial a_i}\right] \\ &+ \left[\frac{\partial u_j(a_j^P(\theta'_J), a_i^P(\theta'_J))}{\partial a_i} - \frac{\partial u_j(a_j^P(\theta_J), a_i^P(\theta_J))}{\partial a_i}\right] = \theta'_J - \theta_J\end{aligned}$$

Because $\theta'_J > \theta_J$, RHS is greater than zero. Now look at LHS. Take the first bracketed term:

$$\frac{\partial u_i(a_i^P(\theta'_J), a_j^P(\theta'_J))}{\partial a_i} - \frac{\partial u_i(a_i^P(\theta_J), a_j^P(\theta_J))}{\partial a_i}$$

$$\begin{aligned}&\left(\frac{\partial u_i(a_i^P(\theta'_J), a_j^P(\theta'_J))}{\partial a_i} - \frac{\partial u_i(a_i^P(\theta_J), a_j^P(\theta'_J))}{\partial a_i}\right) \\ &+ \left(\frac{\partial u_i(a_i^P(\theta_J), a_j^P(\theta'_J))}{\partial a_i} - \frac{\partial u_i(a_i^P(\theta_J), a_j^P(\theta_J))}{\partial a_i}\right)\end{aligned}$$

Since u is concave in own action, if $a_i^P(\theta'_J) > a_i^P(\theta_J)$, then it must be the case that for any a_j:

$$\frac{\partial u_i(a_i^P(\theta'_J), \cdot)}{\partial a_i} < \frac{\partial u_i(a_i^P(\theta_J), \cdot)}{\partial a_i}$$

Furthermore, since u is submodular in opponent action, if

$$a_j^P(\theta'_J) > a_j^P(\theta_J),$$

then it must be the case that for any a_i:

$$\frac{\partial u_i(\cdot, a_j^P(\theta'_J))}{\partial a_i} < \frac{\partial u_i(\cdot, a_j^P(\theta_J))}{\partial a_i}$$

This means that the LHS is negative, so it cannot equal the positive RHS. This is a contradiction, so this case will not occur.

ii. Case ii. $a^P(\theta'_J) = a^P(\theta_J)$

Look at the FOC for $\frac{\partial(w_i+w_j)}{\partial a_i}$ (the FOC for $\frac{\partial(w_i+w_j)}{\partial a_j}$ are symmetric):

$$\frac{\partial u_i(a_i^P(\theta_J), a_j^P(\theta_J))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_J), a_i^P(\theta_J))}{\partial a_i} = \theta_J$$
$$\frac{\partial u_i(a_i^P(\theta'_J), a_j^P(\theta'_J))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta'_J), a_i^P(\theta'_J))}{\partial a_i} = \theta'_J$$

If $a^P(\theta'_J) = a^P(\theta_J)$, that means that the two LHS are equal as well. This implies that the two RHS should be equal, so $\theta_J = \theta'_J$. This is a contradiction of $\theta_J < \theta'_J$, so this case cannot occur.

Since both cases are contradictions, it must be that for $\theta'_J > \theta_J$, then $a_i^P(\theta'_J) < a_i^P(\theta_J)$, so the social planner's chosen action is decreasing in θ_J. $\square$

By Lemma C.1, it is seen that:

$$\frac{\partial a_i^P(\cdot)}{\partial \theta_J} < 0.$$

Combining this result with that of non-coordination, it has been obtained that for all agents i:

$$\frac{d}{d\theta_J}\left[a_i^N(\theta_J) - a_i^P(\theta_J)\right] > 0.$$

2. **Own Effect**

The own effect is set-up in the paper in Equations (2.6) and (2.7). First, I examine the Nash first order conditions, and then I examine the social planner's first order conditions.

(a) Non-coordination: The maximization problem for agent i, given the Nash equilibrium action of agent j, is:

$$\max_{a_i \in A_i} \; u_i(a_i, a_j^N(\theta_I)) + \theta_I a_i$$

$a^N(\theta_I)$ solves the following:

$$\frac{\partial w_i}{\partial a_i} = \frac{\partial u_i(\cdot, a_j^N(\theta_I))}{\partial a_i} + \theta_I \equiv 0$$
$$\frac{\partial w_j}{\partial a_j} = \frac{\partial u_j(\cdot, a_i^N(\theta_I))}{\partial a_j} + \theta_I \equiv 0$$

The intuition is that this action is increasing, due to the increased own benefit.[1]

Lemma C.2. *$a_i^N(\theta_I)$ is increasing in θ_I.*

Proof. Suppose not. Suppose that for $\theta_I' > \theta_I$, $a_i^N(\theta_I') \not> a_i^N(\theta_I)$.

i. Case i. $a_i^N(\theta_I') < a_i^N(\theta_I)$

   Look at the FOC for $\frac{\partial w_i}{\partial a_i}$:

$$\frac{\partial u_i(a_i^N(\theta_I), a_j^N(\theta_I))}{\partial a_i} = -\theta_I$$
$$\frac{\partial u_i(a_i^N(\theta_I'), a_j^N(\theta_I'))}{\partial a_i} = -\theta_I'$$

   Subtract the first from the second:

$$\frac{\partial u_i(a_i^N(\theta_I'), a_j^N(\theta_I'))}{\partial a_i} - \frac{\partial u_i(a_i^N(\theta_I), a_j^N(\theta_I))}{\partial a_i} = -\theta_I' + \theta_I$$

   The RHS is negative. Look at the LHS and add/subtract some terms:

$$\frac{\partial u_i(a_i^N(\theta_I'), a_j^N(\theta_I'))}{\partial a_i} - \frac{\partial u_i(a_i^N(\theta_I), a_j^N(\theta_I))}{\partial a_i}$$
$$\frac{\partial u_i(a_i^N(\theta_I'), a_j^N(\theta_I'))}{\partial a_i} - \frac{\partial u_i(a_i^N(\theta_I), a_j^N(\theta_I'))}{\partial a_i}$$
$$+\frac{\partial u_i(a_i^N(\theta_I), a_j^N(\theta_I'))}{\partial a_i} - \frac{\partial u_i(a_i^N(\theta_I), a_j^N(\theta_I))}{\partial a_i}$$

[1] In the non-symmetric game, there are cross-partials to check to determine the direction of change, but the symmetric game imposes additional assumptions that assist in making this straightforward.

If $a_i^N(\theta_I') < a_i^N(\theta_I)$, then because of concavity in own action, the first subtraction pair is positive. Since the agents are symmetric, agent j's action must follow the same pattern. If $a_j^N(\theta_I') < a_j^N(\theta_I)$, then because of submodularity, the second subtraction pair is also positive. Thus the LHS is positive, which contradicts the RHS being negative. Thus, this case cannot occur.

ii. Case ii. $a_i^N(\theta_I') = a_i^N(\theta_I)$

Look at the FOC for $\frac{\partial w_i}{\partial a_i}$:

$$\frac{\partial u_i(a_i^N(\theta_I), a_j^N(\theta_I))}{\partial a_i} = -\theta_I$$
$$\frac{\partial u_i(a_i^N(\theta_I'), a_j^N(\theta_I'))}{\partial a_i} = -\theta_I'$$

If $a_i^N(\theta_I') = a_i^N(\theta_I)$ and $a_j^N(\theta_I') = a_j^N(\theta_I)$, then the LHS of both of these are equal. This means the RHS should be equal too. This is a contradiction of the assumption that $\theta_I' > \theta_I$. Therefore, this case cannot occur.

Since both cases cannot occur, it must be the case that $a_i^N(\theta)$ is increasing in θ. This holds symmetrically for $a_J^N(\theta)$. □

By Lemma C.2, it is obtained that:

$$\frac{\partial a_i^N(\cdot)}{\partial \theta_I} > 0.$$

(b) Coordination: The maximization problem for the social planner is:

$$\max_{(a_i,a_j)\in A_i\times A_j} u_i(a_i,a_j)+u_j(a_j,a_i)+\theta_I\,(a_i+a_j)$$

For $\theta_I \geq 0$, $a^P(\theta_I)$ solves the following:

$$\frac{\partial(w_i+w_j)}{\partial a_i} = \frac{\partial u_i(\cdot, a_j^P(\theta_I))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_I),\cdot)}{\partial a_i} + \theta_I \equiv 0$$
$$\frac{\partial(w_i+w_j)}{\partial a_j} = \frac{\partial u_j(\cdot, a_i^P(\theta_I))}{\partial a_j} + \frac{\partial u_i(a_i^P(\theta_I),\cdot)}{\partial a_j} + \theta_I \equiv 0$$

The Social Planner may also want to increase actions, because of the increased benefit, but will be wary of the submodularity's effect as well. Recall here, because this is a symmetric game, the action changes will go in the same direction for both agents. The next lemma posits that the SP's actions are also increasing.

Lemma C.3. *$a^P(\theta_I)$ is increasing in θ_I.*

Proof. Suppose not. Suppose that for $\theta_I' > \theta_I$, $a^P(\theta_I') \not> a^P(\theta_I)$.

i. Case i. $a^P(\theta_I') < a^P(\theta_I)$

   Look at the FOC for $\frac{\partial(w_i+w_j)}{\partial a_i}$ (the FOC for $\frac{\partial(w_i+w_j)}{\partial a_j}$ are symmetric):

$$\frac{\partial u_i(a_i^P(\theta_I), a_j^P(\theta_I))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_I), a_i^P(\theta_I))}{\partial a_i} = -\theta_I$$
$$\frac{\partial u_i(a_i^P(\theta_I'), a_j^P(\theta_I'))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_I'), a_i^P(\theta_I'))}{\partial a_i} = -\theta_I'$$

Subtract the second from the first:

$$\frac{\partial u_i(a_i^P(\theta_I'), a_j^P(\theta_I'))}{\partial a_i} - \frac{\partial u_i(a_i^P(\theta_I), a_j^P(\theta_I))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_I'), a_i^P(\theta_I'))}{\partial a_i}$$
$$- \frac{\partial u_j(a_j^P(\theta_I), a_i^P(\theta_I))}{\partial a_i} = -\theta_I' + \theta_I$$

The RHS is negative. Examine the first subtraction pair of the LHS:

$$\frac{\partial u_i(a_i^P(\theta_I'), a_j^P(\theta_I'))}{\partial a_i} - \frac{\partial u_i(a_i^P(\theta_I), a_j^P(\theta_I))}{\partial a_i}$$
$$\left(\frac{\partial u_i(a_i^P(\theta_I'), a_j^P(\theta_I'))}{\partial a_i} - \frac{\partial u_i(a_i^P(\theta_I), a_j^P(\theta_I'))}{\partial a_i}\right)$$
$$+ \left(\frac{\partial u_i(a_i^P(\theta_I), a_j^P(\theta_I'))}{\partial a_i} - \frac{\partial u_i(a_i^P(\theta_I), a_j^P(\theta_I))}{\partial a_i}\right)$$

If $a^P(\theta_I') < a^P(\theta_I)$, then by concavity wrt own action, the first subtraction pair is positive, and by submodularity, the second pair is positive. This holds for agent j's first derivatives as well, so the LHS of the previous statement is positive. This contradicts the negative LHS, so this case cannot occur.

ii. Case ii. $a^P(\theta_I') = a^P(\theta_I)$

Look at the FOC for $\frac{\partial(w_i+w_j)}{\partial a_i}$ (the FOC for $\frac{\partial(w_i+w_j)}{\partial a_j}$ are symmetric):

$$\frac{\partial u_i(a_i^P(\theta_I), a_j^P(\theta_I))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_I), a_i^P(\theta_I))}{\partial a_i} = -\theta_I$$
$$\frac{\partial u_i(a_i^P(\theta_I'), a_j^P(\theta_I'))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_I'), a_i^P(\theta_I'))}{\partial a_i} = -\theta_I'$$

If $a^P(\theta'_I) = a^P(\theta_I)$, then the LHS of both functions must be the same. This means the RHS must be the same, i.e. $\theta'_I = \theta_I$, but this is a contradiction. Therefore, this case cannot occur.

Since both of these cases cannot occur, it must be that a_i^P is increasing in θ_I. $\square$

By Lemma C.3, it is obtained that:

$$\frac{\partial a_i^P(\cdot)}{\partial \theta_I} > 0.$$

Both the Nash actions and the efficient actions are increasing in θ_I. At each θ_I, it should be that the efficient actions are smaller than the Nash actions because of the negative externality. Intuition says that the Nash increases are larger, because the agents ignore the externality, but this really depends on the curvature of the utility function. Therefore, though the directions actions take are known, as is the increase in utility for social planner problem, the ambiguity in utility for the Nash problem makes it difficult to say whether the Nash increases are larger or smaller than the efficient increases, rendering the comparison ambiguous.

Thus, the own effect is the confusing type of externality. One the one hand, the direct benefit increases utility, but on the other hand, the agents then exert more of the externality on each other.

3. **Submodular Effect**

The submodular effect is set-up in the paper in Equations (2.8) and (2.9). First, I examine the Nash first order conditions, and then I examine the social planner's first order conditions.

(a) Non-coordination: The maximization problem for agent i, given the Nash equilibrium action of agent j, is:

$$\max_{a_i \in A_i} \; u_i(a_i, a_j^N(\theta_{IJ})) + \theta_{IJ} a_i a_j^N(\theta_{IJ})$$

$a^N(0)$ solves:

$$\begin{aligned}\frac{\partial w_i}{\partial a_i} &= \frac{\partial u_i\left(\cdot, a_j^N(0)\right)}{\partial a_i} \equiv 0 \\ \frac{\partial w_j}{\partial a_j} &= \frac{\partial u_j\left(\cdot, a_i^N(0)\right)}{\partial a_j} \equiv 0\end{aligned}$$

$a^N(\theta_{IJ})$ solves:

$$\begin{aligned}\frac{\partial w_i}{\partial a_i} &= \frac{\partial u_i\left(\cdot, a_j^N(\theta_{IJ})\right)}{\partial a_i} - \theta_{IJ} a_j^N(\theta_{IJ}) \equiv 0 \\ \frac{\partial w_j}{\partial a_j} &= \frac{\partial u_j\left(\cdot, a_i^N(\theta_{IJ})\right)}{\partial a_j} - \theta_{IJ} a_i^N(\theta_{IJ}) \equiv 0\end{aligned}$$

Going off of the structure above, the next lemma posits that the submodular effect is rendered Null for the Nash equilibrium.

Lemma C.4. *For $\theta'_{IJ} > \theta_{IJ}$, $a_i^N(\theta'_{IJ}) < a_i^N(\theta_{IJ})$.*

Proof. Suppose not. Suppose that for $\theta'_{IJ} > \theta_{IJ}$, $a_i^N(\theta_{IJ}) \geq a_i^N(\theta'_{IJ})$. Because of symmetric utility functions, agents actions go in the same direction.

i. Case i. $a_i^N(\theta'_{IJ}) > a_i^N(\theta_{IJ})\,\forall i$. Look at the FOC for $\frac{\partial w_i}{\partial a_i}$:

$$\frac{\partial u_i\left(a_i^N(\theta_{IJ}), a_j^N(\theta_{IJ})\right)}{\partial a_i} = \theta_{IJ} a_j^N(\theta_{IJ})$$
$$\frac{\partial u_i\left(a_i^N(\theta'_{IJ}), a_j^N(\theta'_{IJ})\right)}{\partial a_i} = \theta'_{IJ} a_j^N(\theta'_{IJ})$$

Subtract the second from the first:

$$\frac{\partial u_i\left(a_i^N(\theta'_{IJ}), a_j^N(\theta'_{IJ})\right)}{\partial a_i} - \frac{\partial u_i\left(a_i^N(\theta_{IJ}), a_j^N(\theta_{IJ})\right)}{\partial a_i} = \theta'_{IJ} a_j^N(\theta'_{IJ}) - \theta_{IJ} a_j^N(\theta_{IJ})$$

RHS is positive. If both Nash actions are larger, then by concavity and submodularity, LHS is negative. This case cannot occur.

ii. Case ii. $a_i^N(\theta'_{IJ}) = a_i^N(\theta_{IJ})\,\forall i$.

$$\frac{\partial u_i\left(a_i^N(\theta'_{IJ}), a_j^N(\theta'_{IJ})\right)}{\partial a_i} - \frac{\partial u_i\left(a_i^N(\theta_{IJ}), a_j^N(\theta_{IJ})\right)}{\partial a_i} = \theta'_{IJ} a_j^N(\theta'_{IJ}) - \theta_{IJ} a_j^N(\theta_{IJ})$$

If actions are equal, then LHS is equal to zero. The statement can be rewritten as:

$$0 = \left(\theta'_{IJ} - \theta_{IJ}\right) a_j^N(\theta_{IJ})$$

In order for RHS to be zero, need $\theta'_{IJ} = \theta_{IJ}$. This is a contradiction.

Therefore the only possibility is that $a_i^N(\theta_{IJ})$ to be decreasing in θ_{IJ}.

$\square$

By Lemma C.4, it is obtained that:

$$\frac{\partial a_i^N(\cdot)}{\partial \theta_{IJ}} < 0.$$

(b) Coordination: The maximization problem for the social planner is:

$$\max_{(a_i,a_j)\in A_i\times A_j} u_i(a_i,a_j) + u_j(a_j,a_i) + 2\theta_{IJ}a_ia_j$$

For $\theta_{IJ} \geq 0$, $a^P(\theta_J)$ solves the following:

$$\frac{\partial(w_i+w_j)}{\partial a_i} = \frac{\partial u_i(\cdot, a_j^P(\theta_{IJ}))}{\partial a_i} + \frac{\partial u_j(a_j^P(\theta_{IJ}),\cdot)}{\partial a_i} - 2\theta_{IJ}a_j^P(\theta_{IJ}) \equiv 0$$
$$\frac{\partial(w_i+w_j)}{\partial a_j} = \frac{\partial u_j(\cdot, a_i^P(\theta_{IJ}))}{\partial a_j} + \frac{\partial u_i(a_i^P(\theta_{IJ}),\cdot)}{\partial a_j} - 2\theta_{IJ}a_i^P(\theta_{IJ}) \equiv 0$$

Lemma C.5. *For $\theta'_{IJ} > \theta_{IJ}$, $a_i^P(\theta'_{IJ}) < a_i^P(\theta_{IJ})$.*

Proof. Suppose not. Suppose $a^P(\theta'_{IJ}) \geq a^P(\theta_{IJ})$.

i. Case i. $a_i^P(\theta'_{IJ}) > a_i^P(\theta_{IJ})\,\forall i$. Look at the FOC for $\frac{\partial w_i}{\partial a_i}$:

$$\frac{\partial u_i\left(a_i^P(\theta_{IJ}), a_j^P(\theta_{IJ})\right)}{\partial a_i} + \frac{\partial u_j\left(a_j^P(\theta_{IJ}), a_i^P(\theta_{IJ})\right)}{\partial a_i} = 2\theta_{IJ}a_j^P(\theta_{IJ})$$
$$\frac{\partial u_i\left(a_i^P(\theta'_{IJ}), a_j^P(\theta'_{IJ})\right)}{\partial a_i} + \frac{\partial u_j\left(a_j^P(\theta'_{IJ}), a_i^P(\theta'_{IJ})\right)}{\partial a_i} = 2\theta'_{IJ}a_j^P(\theta'_{IJ})$$

Subtract the second from the first:

$$
\begin{aligned}
&\frac{\partial u_i\left(a_i^P(\theta'_{IJ}),a_j^P(\theta'_{IJ})\right)}{\partial a_i}+\frac{\partial u_j\left(a_j^P(\theta'_{IJ}),a_i^P(\theta'_{IJ})\right)}{\partial a_i}\\
-&\frac{\partial u_i\left(a_i^P(\theta_{IJ}),a_j^P(\theta_{IJ})\right)}{\partial a_i}-\frac{\partial u_j\left(a_j^P(\theta_{IJ}),a_i^P(\theta_{IJ})\right)}{\partial a_i}\\
=&2\theta'_{IJ}a_j^P(\theta'_{IJ})-2\theta_{IJ}a_j^P(\theta_{IJ})
\end{aligned}
$$

If $a_i^P(\theta_{IJ}) > a_i^P(\theta'_{IJ})$, then RHS is positive.

Because of concavity and submodularity, when both $a_i^P(\theta'_{IJ}) > a_i^P(\theta_{IJ})$ and $a_j^P(\theta'_{IJ}) > a_j^P(\theta_{IJ})$, we have that:

$$\frac{\partial u_i(a_i^P(\theta'_{IJ}),a_j^P(\theta'_{IJ}))}{\partial a_i} < \frac{\partial u_i(a_j^P(\theta_{IJ}),a_j^P(\theta_{IJ}))}{\partial a_i}$$

and that:

$$\frac{\partial u_j(a_i^P(\theta'_{IJ}),a_j^P(\theta'_{IJ}))}{\partial a_i} < \frac{\partial u_j(a_j^P(\theta_{IJ}),a_j^P(\theta_{IJ}))}{\partial a_i}$$

This means that LHS is negative, which is a contradiction. This case cannot occur.

ii. Case ii. $a_i^P(\theta'_{IJ}) = a_i^P(\theta_{IJ})$

If the actions are equal for both agents, then LHS is zero, and the subtracted FOC can be written as:

$$0 = 2\left(\theta'_{IJ}-\theta_{IJ}\right)a_j^P(\theta_{IJ})$$

The only way for RHS to equal LHS is for $\theta'_{IJ} = \theta_{IJ}$. This is a contradiction, so this case cannot occur.

□

By Lemma C.5, it is obtained that:

$$\frac{\partial a_i^P(\cdot)}{\partial \theta_{IJ}} < 0.$$

Both the Nash actions and the efficient actions are decreasing in θ_{IJ}. Similar as with the own effect, the efficient actions should be smaller. From the extra two in the social planner's first order conditions, it is suspected that the efficient actions are decreasing more quickly than the non-coordination actions, but this depends on the curvature of the utility function. Thus, the submodular effect is ambiguous as well.[2]

Combined, these three results give Theorem 2.1.

□

C.3 Expanded Results from Section 2.3.2

<u>Derivation of Second-Order Expansions</u>

1. Non-coordination first order conditions:

   Recall the FOC are:

$$\frac{\partial u_i(a_i, a_j)}{\partial a_i} + \theta_{ii} \equiv 0$$
$$\frac{\partial u_j(a_j, a_i)}{\partial a_j} + \theta_{jj} \equiv 0$$

[2]For a centered Taylor expansion version of this problem, the submodular effect would be null, as opposed to ambiguous.

The derivatives of these with respect to agent i's own effect, θ_{ii}, are:

$$\begin{aligned}
\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2}\cdot\frac{\partial a_i^N(\theta)}{\partial\theta_{ii}}+\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i\partial a_j}\cdot\frac{\partial a_j^N(\theta)}{\partial\theta_{ii}}+1 &= 0\\
\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j\partial a_i}\cdot\frac{\partial a_i^N(\theta)}{\partial\theta_{ii}}+\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j^2}\cdot\frac{\partial a_j^N(\theta)}{\partial\theta_{ii}} &= 0
\end{aligned}$$

and with respect to agent i's opponent effect, θ_{ij}, are:

$$\begin{aligned}
\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2}\cdot\frac{\partial a_i^N(\theta)}{\partial\theta_{ij}}+\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i\partial a_j}\cdot\frac{\partial a_j^N(\theta)}{\partial\theta_{ij}} &= 0\\
\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j\partial a_i}\cdot\frac{\partial a_i^N(\theta)}{\partial\theta_{ij}}+\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j^2}\cdot\frac{\partial a_j^N(\theta)}{\partial\theta_{ij}} &= 0
\end{aligned}$$

In the second set of expansions, those with respect to θ_{ij}, since the function is concave and submodular, then in both top and bottom two negative numbers multiplied by the derivatives. In order for any set of numbers other than zero to solve this set of equations, this would require the own second derivatives to equal the cross-partials, which is possible, but a small set of functions. Furthermore, since the parameter θ_{ij} does not appear in the first order conditions, this proof will proceed with the case of:

$$\frac{\partial a_i^N(\theta)}{\partial\theta_{ij}}=\frac{\partial a_j^N(\theta)}{\partial\theta_{ij}}=0.$$

The derivatives of the FOC with respect to agent j's opponent effect on i, θ_{ji}, are:

$$\begin{aligned}
\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2}\cdot\frac{\partial a_i^N(\theta)}{\partial\theta_{ji}}+\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i\partial a_j}\cdot\frac{\partial a_j^N(\theta)}{\partial\theta_{ji}} &= 0\\
\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j\partial a_i}\cdot\frac{\partial a_i^N(\theta)}{\partial\theta_{ji}}+\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j^2}\cdot\frac{\partial a_j^N(\theta)}{\partial\theta_{ji}} &= 0
\end{aligned}$$

and with respect to agent j's own effect , θ_{jj}, are:

$$\begin{aligned}
\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2}\cdot\frac{\partial a_i^N(\theta)}{\partial \theta_{jj}}+\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j}\cdot\frac{\partial a_j^N(\theta)}{\partial \theta_{jj}} \qquad &= 0\\
\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j \partial a_i}\cdot\frac{\partial a_i^N(\theta)}{\partial \theta_{jj}}+\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j^2}\cdot\frac{\partial a_j^N(\theta)}{\partial \theta_{jj}}+1 &= 0
\end{aligned}$$

The parameter θ_{ji} displays a similar pattern as did θ_{ij}, and so the proof will proceed under the following:

$$\frac{\partial a_i^N(\theta)}{\partial \theta_{ji}}=\frac{\partial a_j^N(\theta)}{\partial \theta_{ji}}=0.$$

The results with respect to θ_{ii} and θ_{jj} can be condensed into matrix form:

$$\begin{bmatrix}\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2} & \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j}\\ \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j} & \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2}\end{bmatrix}\cdot\begin{bmatrix}\frac{\partial a_i^N}{\partial \theta_{ii}} & \frac{\partial a_i^N}{\partial \theta_{jj}}\\ \frac{\partial a_j^N}{\partial \theta_{ii}} & \frac{\partial a_j^N}{\partial \theta_{jj}}\end{bmatrix}=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}$$

2. Coordination first order conditions:

   Recall the FOC are:

$$\frac{\partial u_i(a_i,a_j)}{\partial a_i}+\frac{\partial u_i(a_i,a_j)}{\partial a_j}+\frac{\partial u_j(a_j,a_i)}{\partial a_i}+\frac{\partial u_j(a_j,a_i)}{\partial a_j}+\gamma_i+\gamma_j\equiv 0$$

   The derivatives of these with respect to agent i's total effect, γ_i, are:

$$\begin{aligned}
&\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2}\cdot\frac{\partial a_i^P}{\partial \gamma_i}+\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j}\cdot\frac{\partial a_j^P}{\partial \gamma_i}\\
&\qquad +\frac{\partial^2 u_j(a_j,a_i)}{\partial a_i^2}\cdot\frac{\partial a_i^P}{\partial \gamma_i}+\frac{\partial^2 u_j(a_j,a_i)}{\partial a_i \partial a_j}\cdot\frac{\partial a_j^P}{\partial \gamma_i}+1 \quad = 0\\
&\frac{\partial^2 u_i(a_i,a_j)}{\partial a_j \partial a_i}\cdot\frac{\partial a_i^P}{\partial \gamma_i}+\frac{\partial^2 u_i(a_i,a_j)}{\partial a_j^2}\cdot\frac{\partial a_j^P}{\partial \gamma_i}\\
&\qquad +\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j \partial a_i}\cdot\frac{\partial a_i^P}{\partial \gamma_i}+\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j^2}\cdot\frac{\partial a_j^P}{\partial \gamma_i} \qquad = 0
\end{aligned}$$

and the derivatives with respect to agent j's total effect, γ_j, are:

$$\begin{aligned}
&\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2}\cdot\frac{\partial a_i^P}{\partial \gamma_j}+\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j}\cdot\frac{\partial a_j^P}{\partial \gamma_j} \\
&\qquad +\frac{\partial^2 u_j(a_j,a_i)}{\partial a_i^2}\cdot\frac{\partial a_i^P}{\partial \gamma_j}+\frac{\partial^2 u_j(a_j,a_i)}{\partial a_i \partial a_j}\cdot\frac{\partial a_j^P}{\partial \gamma_j} &&= 0 \\
&\frac{\partial^2 u_i(a_i,a_j)}{\partial a_j \partial a_i}\cdot\frac{\partial a_i^P}{\partial \gamma_j}+\frac{\partial^2 u_i(a_i,a_j)}{\partial a_j^2}\cdot\frac{\partial a_j^P}{\partial \gamma_j} \\
&\qquad +\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j \partial a_j}\cdot\frac{\partial a_i^P}{\partial \gamma_j}+\frac{\partial^2 u_j(a_j,a_i)}{\partial a_j^2}\cdot\frac{\partial a_j^P}{\partial \gamma_j}+1 &&= 0
\end{aligned}$$

The results with respect to γ_i and γ_j can be condensed into matrix form:

$$\begin{bmatrix}
\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2}+\frac{\partial^2 u_j(a_i,a_j)}{\partial a_i^2} & \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j}+\frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j} \\
\frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j}+\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} & \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2}+\frac{\partial^2 u_i(a_i,a_j)}{\partial a_j^2}
\end{bmatrix}\cdot\begin{bmatrix}
\frac{\partial a_i^P}{\partial \gamma_i} & \frac{\partial a_i^P}{\partial \gamma_j} \\
\frac{\partial a_j^P}{\partial \gamma_i} & \frac{\partial a_j^P}{\partial \gamma_j}
\end{bmatrix}=\begin{bmatrix}
-1 & 0 \\
0 & -1
\end{bmatrix}$$

Proof of Lemma 2.2

Restatement of Lemma 2.2. *For Da^N to be positive and for the Nash actions to be increasing in response to an increase in θ, it is sufficient for the own second derivatives to be the same direction in comparison to the cross-partials for both agents. That is, the own second derivative can be more negative than the cross partial for both agents:*

$$\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2}<\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} \quad \text{and} \quad \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2}<\frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j}$$

or, the own second derivative can be less negative than the cross partial for both agents:

$$\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2}>\frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} \quad \text{and} \quad \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2}>\frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j}$$

Proof. Recall the set-up of the linearization:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2} & \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} \\ \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j} & \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2} \end{bmatrix}$$

$$ad - bc = \frac{\partial^2 u_i(a_i, a_j)}{\partial a_i^2} \cdot \frac{\partial^2 u_j(a_i, a_j)}{\partial a_j^2} - \frac{\partial^2 u_i(a_i, a_j)}{\partial a_i \partial a_j} \cdot \frac{\partial^2 u_j(a_i, a_j)}{\partial a_i \partial a_j}$$

$$\begin{bmatrix} d - c & a - b \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 u_j(a_i,a_j)}{\partial a_j^2} - \frac{\partial^2 u_j(a_i,a_j)}{\partial a_i \partial a_j} & \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i^2} - \frac{\partial^2 u_i(a_i,a_j)}{\partial a_i \partial a_j} \end{bmatrix}$$

There are two cases to consider:

1. $ad - bc > 0$ while $d - c < 0, a - b < 0$

$$ad - bc = \frac{\partial^2 u_i(a_i, a_j)}{\partial a_i^2} \cdot \frac{\partial^2 u_j(a_i, a_j)}{\partial a_j^2} - \frac{\partial^2 u_i(a_i, a_j)}{\partial a_i \partial a_j} \cdot \frac{\partial^2 u_j(a_i, a_j)}{\partial a_i \partial a_j}$$

$$ad - bc > 0 \Rightarrow$$

$$\frac{\partial^2 u_i(a_i, a_j)}{\partial a_i^2} \cdot \frac{\partial^2 u_j(a_i, a_j)}{\partial a_j^2} > \frac{\partial^2 u_i(a_i, a_j)}{\partial a_i \partial a_j} \cdot \frac{\partial^2 u_j(a_i, a_j)}{\partial a_i \partial a_j}$$

By concavity and submodularity, these are individually negative. So, the above statement could hold under some sort of dominant effect idea, where the own second derivative is more negative ("larger") than the cross partial. Then $d - c$ and $a - b$ would be negative, because both d and a would be smaller (more negative) than c and b.

2. $ad - bc < 0$ and $d - c > 0, a - b > 0$

On the other hand, with the opposite of dominant effect, or some second order opponent effect, then ad would be smaller than bc, but c would be smaller from d (as well as b from a, which would be positive). This would give the same required sign.

$\square$

<u>Proof of Lemma 2.3</u>

Restatement of Lemma 2.3. *For U^P to have only positive eigenvalues, it is sufficient that:*

$$\frac{\partial^2 u_i}{\partial a_j^2} > 0$$

$$\frac{\partial^2 u_j}{\partial a_j^2} > 0$$

$$\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) > \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2$$

Proof. Recall the method for calculating eigenvalues:

$$\begin{bmatrix} \frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} - \lambda & \frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j} \\ \frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j} & \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2} - \lambda \end{bmatrix}$$

The characteristic function is:

$$\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} - \lambda\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2} - \lambda\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2 = 0$$

Expanding:

$$\lambda^2 - \lambda\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) + \left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2 = 0$$

Using quadratic function, we know that the values for λ are as follows:

$$\lambda = \frac{1}{2}\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) \pm \sqrt{\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)^2 - 4\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2\right]}\right]$$

When are both λ positive? First, check under the square root.

$$\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)^2 \overset{?}{>} 4\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2\right]$$

It is known that the eigenvectors of symmetric matrices are real, so the above equation must hold. Therefore, what is under the square root must be positive. Hence, the following condition that must be also true:

$$\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)^2 + \left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)^2 + \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2 > 2\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)$$

The two eigenvalues can be denoted as:

$$\lambda_1 = \frac{1}{2}\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) + \sqrt{\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)^2 - 4\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2\right]}\right]$$

$$\lambda_2 = \frac{1}{2}\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \sqrt{\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)^2 - 4\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2\right]}\right]$$

From this, it is is clear that if $\lambda_2 > 0 \Rightarrow \lambda_1 > 0$ (adding a positive amount vs. subtracting it). Hence, for both to be positive, the minimum is to check when λ_2 is positive.

$$0 < \frac{1}{2}\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \sqrt{\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)^2 - 4\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2\right]}\right]$$

$$0 < \left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \sqrt{\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)^2 - 4\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2\right]}$$

$$\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) > \sqrt{\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)^2 - 4\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2\right]}$$

The square root must be positive, so that means

$$\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) > 0$$

as well, and since the utility function is concave, need $\frac{\partial^2 u_i}{\partial a_j^2} > 0$ and $\frac{\partial^2 u_j}{\partial a_j^2} > 0$. With those added assumptions, square both sides:

$$\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)^2 > \left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2} + \frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right)^2$$
$$-4\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2\right]$$

This then becomes:

$$0 > -4\left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2\right]$$
$$0 < \left[\left(\frac{\partial^2 u_i}{\partial a_i^2} + \frac{\partial^2 u_j}{\partial a_i^2}\right)\left(\frac{\partial^2 u_i}{\partial a_j^2} + \frac{\partial^2 u_j}{\partial a_j^2}\right) - \left(\frac{\partial^2 u_i}{\partial a_i \partial a_j} + \frac{\partial^2 u_j}{\partial a_i \partial a_j}\right)^2\right]$$

Hence, convex opponent derivative and the above condition are sufficient for positive eigenvalues. $\square$

Appendix D

Proofs and Derivations for Chapter 3

D.1 Proofs for Fallow Fisheries from Section 3.3

D.1.1 Derivations and Proofs under Non-coordination

Finding $\Pi_g^i(p_i, p_j^M)$

$$\begin{aligned}
\Pi_g^i(p_i, p_j^M) &= u_1 + \beta\left[(1-r)\Pi_g^i(p_i, p_j^M) + r\Pi_b^i(p_i, p_j^M)\right] \\
(1-\beta(1-r))\,\Pi_g^i(p_i, p_j^M) &= u_1 + \beta r \Pi_b^i(p_i, p_j^M) \\
\Pi_g^i(p_i, p_j^M) &= \frac{u_1 + \beta r \Pi_b^i(p_i, p_j^M)}{(1-\beta(1-r))}
\end{aligned}$$

Finding $\Pi_b^i(p_i, p_j^M)$

$$\Pi_b^i(p_i, p_j^M) = u_0(p_i) + \beta\left[(p_i + p_j^M)\left(\frac{u_1 + \beta r \Pi_b^i(p_i, p_j^M)}{(1-\beta(1-r))}\right) + (1 - p_i - p_j^M)\Pi_b^i(p_i, p_j^M)\right]$$

$$\left(1-\beta(1-p_i-p_j^M)\right)\Pi_b^i(p_i, p_j^M) = u_0(p_i) + \beta(p_i + p_j^M)\left(\frac{u_1 + \beta r \Pi_b^i(p_i, p_j^M)}{(1-\beta(1-r))}\right)$$

Multiply through by $(1-\beta(1-r))$:

$$\begin{aligned}
&\left(1-\beta(1-p_i-p_j^M)\right)(1-\beta(1-r))\,\Pi_b^i(p_i, p_j^M) \\
&\qquad = (1-\beta(1-r))\,u_0(p_i) + \beta(p_i+p_j^M)\left(u_1 + \beta r \Pi_b^i(p_i, p_j^M)\right) \\
&\left[\left(1-\beta(1-p_i-p_j^M)\right)(1-\beta(1-r)) - \beta^2 r\left(p_i+p_j^M\right)\right]\Pi_b^i(p_i, p_j^M) \\
&\qquad = (1-\beta(1-r))\,u_0(p_i) + \beta(p_i+p_j^M)u_1
\end{aligned}$$

Simplifying the coefficient on Π_b^i on the LHS:

$$
\begin{aligned}
&\left(1-\beta(1-p_i-p_j^M)\right)\left(1-\beta(1-r)\right)-\beta^2 r\left(p_i+p_j^M\right)\\
=&\,1-\beta(1-r)-\beta(1-p_i-p_j^M)+\beta^2(1-r)(1-p_i-p_j^M)-\beta^2 r\left(p_i+p_j^M\right)\\
=&\,1-\beta(1-r)-\beta(1-p_i-p_j^M)+\beta^2(1-r)(1-\left(p_i+p_j^M\right))-\beta^2 r\left(p_i+p_j^M\right)\\
=&\,1-\beta(1-r)-\beta(1-p_i-p_j^M)+\beta^2(1-\left(p_i+p_j^M\right)-r+r\left(p_i+p_j^M\right))\\
&-\beta^2 r\left(p_i+p_j^M\right)\\
=&\,1-\beta(1-r)-\beta(1-p_i-p_j^M)+\beta^2-\beta^2\left(p_i+p_j^M\right)-\beta^2 r\\
=&\,1-\beta(1-r)-\beta(1-p_i-p_j^M)+\beta^2\left(1-r-p_i-p_j^M\right)
\end{aligned}
$$

Continuing to simplify:

$$
\begin{aligned}
&1-\beta(1-r)-\beta(1-p_i-p_j^M)+\beta^2\left(1-r-p_i-p_j^M\right)\\
=&\,1-2\beta+\beta^2+(\beta-\beta^2)r+(\beta-\beta^2)\left(p_i+p_j^M\right)\\
=&\,(1-\beta)^2+\beta(1-\beta)r+\beta(1-\beta)\left(p_i+p_j^M\right)\\
=&\,(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]\Pi_b^i(p_i,p_j^M)&=\left(1-\beta(1-r)\right)u_0(p_i)+\beta(p_i+p_j^M)u_1\\
\Pi_b^i(p_i,p_j^M)&=\frac{\left(1-\beta(1-r)\right)u_0(p_i)+\beta(p_i+p_j^M)u_1}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]}
\end{aligned}
$$

<u>Proof of Lemma 3.1</u>

Restatement of Lemma 3.1. *There exists a unique interior Markov perfect equilibrium described by:*

$$
\beta u_0(p_i^M)-\left[1-\beta+\beta r+\beta\left(p_i^M+p_j^M\right)\right]u_0'(p_i^M)-\beta u_1=0
$$

$$
\beta u_0(p_j^M)-\left[1-\beta+\beta r+\beta\left(p_i^M+p_j^M\right)\right]u_0'(p_j^M)-\beta u_1=0
$$

Proof. Uniqueness is given by Kuhn-Tucker sufficiency. The function to be maximized is concave in p, and the constraints are linear, so they are convex.

Recall, the maximization problem is:

$$\begin{aligned}\max_{p_i}\ & \frac{(1-\beta(1-r))\,u_0(p_i)+\beta(p_i+p_j^M)u_1}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]}\\ \text{s.t. }\ & p_i \geq \underline{p}\\ & p_i \leq \bar{p}\end{aligned}$$

Assign $+\lambda_{1i}$ to the inequality $p_i-\underline{p}\geq 0$ and $+\lambda_{2i}$ to the inequality $\bar{p}-p_i\geq 0$. Then the Lagrangian is:

$$\mathcal{L}\left(p,\lambda_{1i},\lambda_{2i}\right)=\frac{(1-\beta(1-r))\,u_0(p_i)+\beta(p_i+p_j^M)u_1}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]}+\lambda_{1i}\left(p_i-\underline{p}\right)+\lambda_{2i}\left(\bar{p}-p_i\right)$$

FOC:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial p_i}=&\frac{1}{(1-\beta)^2\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]^2}\cdot\Big\{\left[(1-\beta(1-r))u_0'(p_i)+\beta u_1\right]\\ &\cdot(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]-\beta(1-\beta)\big[\,(1-\beta(1-r))\,u_0(p_i)\\ &+\beta(p_i+p_j^M)u_1\big]\Big\}+\lambda_{1i}-\lambda_{2i}=0\\ \lambda_{1i}\left(p_i-\underline{p}\right)=&0,\lambda_{1i}\geq 0\left(p_i-\underline{p}\right)\geq 0\\ \lambda_{2i}\left(\bar{p}-p_i\right)=&0,\lambda_{2i}\geq 0\left(\bar{p}-p_i\right)\geq 0\end{aligned}$$

The function $\frac{\partial\mathcal{L}}{\partial p_i}$ can be simplified a little bit, since the denominator can cancel with some parts of the numerator. Rewritten as:

$$\frac{\partial \mathcal{L}}{\partial p_i} = \frac{[(1-\beta(1-r))u_0'(p_i) + \beta u_1]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]} - \frac{\beta\left[(1-\beta(1-r))\,u_0(p_i) + \beta(p_i+p_j^M)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]^2} + \lambda_{1i} - \lambda_{2i} = 0$$

The conditions all together are:

$$\frac{\partial \mathcal{L}}{\partial p_i} = \frac{[(1-\beta(1-r))u_0'(p_i) + \beta u_1]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]} - \frac{\beta\left[(1-\beta(1-r))\,u_0(p_i) + \beta(p_i+p_j^M)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]^2} + \lambda_{1i} - \lambda_{2i} = 0$$

$$\lambda_{1i}\left(p_i - \underline{p}\right) = 0, \lambda_{1i} \geq 0\left(p_i - \underline{p}\right) \geq 0$$

$$\lambda_{2i}\left(\bar{p} - p_i\right) = 0, \lambda_{2i} \geq 0\left(\bar{p} - p_i\right) \geq 0$$

Solving, there are three cases to consider:

Case i. $p_i = \underline{p} \Rightarrow \lambda_{2i} = 0, \lambda_{1i} \geq 0$

$$\frac{\left[(1-\beta(1-r))u_0'(\underline{p}) + \beta u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(\underline{p}+p_j^M\right)\right]} + \lambda_{1i} = \frac{\beta\left[(1-\beta(1-r))\,u_0(\underline{p}) + \beta(\underline{p}+p_j^M)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(\underline{p}+p_j^M\right)\right]^2}$$

$$\frac{\left[(1-\beta(1-r))u_0'(\underline{p}) + \beta u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(\underline{p}+p_j^M\right)\right]} \leq \frac{\beta\left[(1-\beta(1-r))\,u_0(\underline{p}) + \beta(\underline{p}+p_j^M)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(\underline{p}+p_j^M\right)\right]^2}$$

Case ii. $p_i = \bar{p} \Rightarrow \lambda_{1i} = 0, \lambda_{2i} \geq 0$

$$\frac{[(1-\beta(1-r))u_0'(\bar{p})+\beta u_1]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(\bar{p}+p_j^M\right)\right]} = \frac{\beta\left[(1-\beta(1-r))\,u_0(\bar{p})+\beta(\bar{p}+p_j^M)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(\bar{p}+p_j^M\right)\right]^2} + \lambda_{2i}$$

$$\frac{[(1-\beta(1-r))u_0'(\bar{p})+\beta u_1]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(\bar{p}+p_j^M\right)\right]} \geq \frac{\beta\left[(1-\beta(1-r))\,u_0(\bar{p})+\beta(\bar{p}+p_j^M)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(\bar{p}+p_j^M\right)\right]^2}$$

Case iii. $\underline{p} < p_i < \bar{p} \Rightarrow \lambda_{1i} = 0, \lambda_{2i} = 0$

$$\frac{[(1-\beta(1-r))u_0'(p_i)+\beta u_1]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]} = \frac{\beta\left[(1-\beta(1-r))\,u_0(p_i)+\beta(p_i+p_j^M)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]^2}$$

$$\left[(1-\beta(1-r))u_0'(p_i)+\beta u_1\right] = \frac{\beta\left[(1-\beta(1-r))\,u_0(p_i)+\beta(p_i+p_j^M)u_1\right]}{\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]}$$

$$\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right][(1-\beta(1-r))u_0'(p_i)+\beta u_1] = \beta\left[(1-\beta(1-r))\,u_0(p_i)+\beta(p_i+p_j^M)u_1\right]$$

$$\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right](1-\beta(1-r))u_0'(p_i) + \beta\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]u_1 = \beta\left(1-\beta(1-r)\right)u_0(p_i)+\beta^2(p_i+p_j^M)u_1$$

$$\left(1-\beta(1-r)\right)\left[\beta u_0(p_i)-\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]u_0'(p_i)\right] = \left[\beta\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]-\beta^2(p_i+p_j^M)\right]u_1$$

Simplifying the coefficient on u_1:

$$
\begin{aligned}
&\beta\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]-\beta^2(p_i+p_j^M)\\
=&\beta-\beta^2+\beta^2 r+\beta^2\left(p_i+p_j^M\right)-\beta^2(p_i+p_j^M)\\
=&\beta-\beta^2+\beta^2 r\\
=&\beta(1-\beta(1-r))
\end{aligned}
$$

$$
\begin{aligned}
(1-\beta(1-r))&\left[\beta u_0(p_i)-\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]u_0'(p_i)\right]\\
&=\beta(1-\beta(1-r))u_1\\
(1-\beta(1-r))&\left[\beta u_0(p_i)-\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]u_0'(p_i)\right]\\
&-\beta(1-\beta(1-r))u_1=0
\end{aligned}
$$

Divide through by $(1-\beta(1-r))\neq 0$:

$$
\beta u_0(p_i)-\left[1-\beta+\beta r+\beta\left(p_i+p_j^M\right)\right]u_0'(p_i)-\beta u_1=0
$$

□

D.1.2 Derivations and Proofs under Coordination

Finding $\Pi_g^P(p_i,p_j)$

$$
\begin{aligned}
\Pi_g^P(p_i,p_j)&=2u_1+\beta\left[(1-r)\Pi_g^P(p_i,p_j)+r\Pi_b^P(p_i,p_j)\right]\\
(1-\beta(1-r))\,\Pi_g^P(p_i,p_j)&=2u_1+\beta r\Pi_b^P(p_i,p_j)\\
\Pi_g^P(p_i,p_j)&=\frac{2u_1+\beta r\Pi_b^P(p_i,p_j)}{(1-\beta(1-r))}
\end{aligned}
$$

Finding $\Pi_b^P(p_i,p_j)$

$$\Pi_b^P(p_i,p_j) = u_0(p_i) + u_0(p_j) + \beta\left[(p_i+p_j)\left(\frac{2u_1 + \beta r \Pi_b^P(p_i,p_j)}{(1-\beta(1-r))}\right) + (1-p_i-p_j)\Pi_b^P(p_i,p_j)\right]$$

$$[1-\beta(1-p_i-p_j)]\,\Pi_b^P(p_i,p_j) = u_0(p_i) + u_0(p_j) + \beta(p_i+p_j)\left(\frac{2u_1 + \beta r \Pi_b^P(p_i,p_j)}{(1-\beta(1-r))}\right)$$

Multiply through by $(1-\beta(1-r))$:

$$\begin{aligned}
&(1-\beta(1-r))\,[1-\beta(1-p_i-p_j)]\,\Pi_b^P(p_i,p_j) \\
&\qquad = (1-\beta(1-r))\,(u_0(p_i)+u_0(p_j)) + \beta(p_i+p_j) \\
&\qquad \cdot \left(2u_1 + \beta r \Pi_b^P(p_i,p_j)\right) \\
&\big[\,(1-\beta(1-r))\,[1-\beta(1-p_i-p_j)] - \beta^2 r(p_i+p_j)\big]\,\Pi_b^P(p_i,p_j) \\
&\qquad = (1-\beta(1-r))\,(u_0(p_i)+u_0(p_j)) + 2\beta(p_i+p_j)u_1
\end{aligned}$$

Using the coefficient simplifications derived for the non-coordination case, this can be rewritten as:

$$(1-\beta)\,[1-\beta+\beta r+\beta\,(p_i+p_j)]\,\Pi_b^P(p_i,p_j) = (1-\beta(1-r))\,(u_0(p_i)+u_0(p_j)) + 2\beta(p_i+p_j)u_1$$

$$\Pi_b^P(p_i,p_j) = \frac{(1-\beta(1-r))\,(u_0(p_i)+u_0(p_j)) + 2\beta(p_i+p_j)u_1}{(1-\beta)\,[1-\beta+\beta r+\beta\,(p_i+p_j)]}$$

Proof of Lemma 3.2

Restatement of Lemma 3.2. *There exists a unique interior Markov perfect cooperative equilibrium given by:*

$$\beta\left[u_0(p_i)+u_0(p_j)\right]-\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]u_0'(p_i)-2\beta u_1=0$$

$$\beta\left[u_0(p_i)+u_0(p_j)\right]-\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]u_0'(p_j)-2\beta u_1=0$$

Proof. Uniqueness is the same as for Lemma 3.1. The first order conditions are:

$$\begin{aligned}\max_{p_i,p_j}\quad & \frac{(1-\beta(1-r))\left(u_0(p_i)+u_0(p_j)\right)+2\beta(p_i+p_j)u_1}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]}\\ \text{s.t.}\quad & p_i\geq\underline{p}\\ & p_i\leq\bar{p}\\ & p_j\geq\underline{p}\\ & p_j\leq\bar{p}\end{aligned}$$

Assign $+\mu_{1i}$ to the inequality $p_i-\underline{p}\geq 0$ and $+\mu_{2i}$ to the inequality $\bar{p}-p_i\geq 0$. Then the Lagrangian is:

$$\begin{aligned}\mathcal{L}\left(p,\mu_{1i},\mu_{2i},\mu_{1j},\mu_{2j}\right)=&\frac{(1-\beta(1-r))\left(u_0(p_i)+u_0(p_j)\right)+2\beta(p_i+p_j)u_1}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]}\\ &+\mu_{1i}\left(p_i-\underline{p}\right)+\mu_{2i}\left(\bar{p}-p_i\right)+\mu_{1j}\left(p_j-\underline{p}\right)+\mu_{2j}\left(\bar{p}-p_j\right)\end{aligned}$$

FOC:

$$\frac{\partial \mathcal{L}}{\partial p_i} = \frac{[(1-\beta(1-r))u_0'(p_i) + 2\beta u_1]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]}$$
$$- \frac{\beta\left[(1-\beta(1-r))\left(u_0(p_i)+u_0(p_j)\right)+2\beta(p_i+p_j)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]^2}$$
$$+ \mu_{1i} - \mu_{2i} = 0$$
$$\frac{\partial \mathcal{L}}{\partial p_j} = \frac{[(1-\beta(1-r))u_0'(p_j) + 2\beta u_1]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]}$$
$$- \frac{\beta\left[(1-\beta(1-r))\left(u_0(p_i)+u_0(p_j)\right)+2\beta(p_i+p_j)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]^2}$$
$$+ \mu_{1j} - \mu_{2j} = 0$$

$$\mu_{1i}\left(p_i-\underline{p}\right)=0, \mu_{1i}\geq 0 \left(p_i-\underline{p}\right)\geq 0$$

$$\mu_{2i}\left(\bar{p}-p_i\right)=0, \mu_{2i}\geq 0 \left(\bar{p}-p_i\right)\geq 0$$

$$\mu_{1j}\left(p_j-\underline{p}\right)=0, \mu_{1j}\geq 0 \left(p_j-\underline{p}\right)\geq 0$$

$$\mu_{2j}\left(\bar{p}-p_j\right)=0, \mu_{2j}\geq 0 \left(\bar{p}-p_j\right)\geq 0$$

There are multiple cases to consider, basically every possible combination of how the multipliers could be.[1] I will list them, but I will only investigate the fully interior one for now:

Case i. $p_i = \underline{p} \Rightarrow \mu_{2i} = 0, \mu_{1i} \geq 0$

Case i.a. $p_j = \underline{p} \Rightarrow \mu_{2j} = 0, \mu_{1j} \geq 0$

Case i.b. $p_j = \bar{p} \Rightarrow \mu_{1j} = 0, \mu_{2j} \geq 0$

Case i.c. $\underline{p} < p_j < \bar{p} \Rightarrow \mu_{1j} = \mu_{2j} = 0$

[1]Only if the agents are somehow non-symmetric, of course.

Case ii. $p_i = \bar{p} \Rightarrow \mu_{1i} = 0, \mu_{2i} \geq 0$

Case ii.a. $p_j = \underline{p} \Rightarrow \mu_{2j} = 0, \mu_{1j} \geq 0$

Case ii.b. $p_j = \bar{p} \Rightarrow \mu_{1j} = 0, \mu_{2j} \geq 0$

Case ii.c. $\underline{p} < p_j < \bar{p} \Rightarrow \mu_{1j} = \mu_{2j} = 0$

Case iii. $\underline{p} < p_i < \bar{p} \Rightarrow \mu_{1i} = \mu_{2i} = 0$

Case iii.a. $p_j = \underline{p} \Rightarrow \mu_{2j} = 0, \mu_{1j} \geq 0$

Case iii.b. $p_j = \bar{p} \Rightarrow \mu_{1j} = 0, \mu_{2j} \geq 0$

Case iii.c. $\underline{p} < p_j < \bar{p} \Rightarrow \mu_{1j} = \mu_{2j} = 0$

This is the purely interior case, and it is the one under consideration. Observe that once the multipliers are gone, it is very clear that these agents are symmetric, so their FOC are quite similar:

$$\frac{[(1-\beta(1-r))u_0'(p_i) + 2\beta u_1]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]} = \frac{\beta\left[(1-\beta(1-r))\left(u_0(p_i)+u_0(p_j)\right)+2\beta(p_i+p_j)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]^2}$$
$$\frac{[(1-\beta(1-r))u_0'(p_j) + 2\beta u_1]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]} = \frac{\beta\left[(1-\beta(1-r))\left(u_0(p_i)+u_0(p_j)\right)+2\beta(p_i+p_j)u_1\right]}{(1-\beta)\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]^2}$$

Multiplying both sides by the LHS denominator:

$$\left[(1-\beta(1-r))u_0'(p_i) + 2\beta u_1\right] = \frac{\beta\left[(1-\beta(1-r))\left(u_0(p_i)+u_0(p_j)\right)+2\beta(p_i+p_j)u_1\right]}{\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]}$$
$$\left[(1-\beta(1-r))u_0'(p_j) + 2\beta u_1\right] = \frac{\beta\left[(1-\beta(1-r))\left(u_0(p_i)+u_0(p_j)\right)+2\beta(p_i+p_j)u_1\right]}{\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]}$$

$$\beta\left[u_0(p_i)+u_0(p_j)\right]-\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]u_0'(p_i)-2\beta u_1=0$$

$$\beta\left[u_0(p_i)+u_0(p_j)\right]-\left[1-\beta+\beta r+\beta\left(p_i+p_j\right)\right]u_0'(p_j)-2\beta u_1=0$$

□

D.1.3 Comparative Statics

Proof of Lemma 3.3

Restatement of Lemma 3.3. *Under the assumptions listed, the social planner's symmetric action p^P is larger than the non-cooperative action p^M.*

Proof. Recall the new assumptions:

$$u_1 > 2u_0(\underline{p})$$

$$0 \geq u_0''(p)$$

$$[1-\beta+\beta r+2\beta p^P] > 0$$

Recall the symmetric conditions for non-coordination and coordination:

$$\beta\left(u_0(p^M)-u_1\right)=\left[1-\beta+\beta r+2\beta p^M\right]u_0'(p^M)$$

$$2\beta\left(u_0(p^P)-u_1\right)=\left[1-\beta+\beta r+2\beta p^P\right]u_0'(p^P)$$

Subtracting from each other:

$$\begin{aligned}
\beta\left[\left(u_0(p^M)-u_1\right)-2\left(u_0(p^P)-u_1\right)\right] &= \left[1-\beta+\beta r+2\beta p^M\right]u_0'(p^M)\\
&\quad-\left[1-\beta+\beta r+2\beta p^P\right]u_0'(p^P)\\
\beta\left[u_0(p^M)-2u_0(p^P)+u_1\right] &= \left[1-\beta+\beta r+2\beta p^M\right]u_0'(p^M)\\
&\quad-\left[1-\beta+\beta r+2\beta p^P\right]u_0'(p^P)
\end{aligned}$$

Suppose not. Suppose $p^P \leq p^M$.

i. Case i. $p^P = p^M$

Let the value be $\mathbf{p}$.

$$\begin{aligned}
\beta\left[u_0(\mathbf{p}) - 2u_0(\mathbf{p}) + u_1\right] &= \left[1 - \beta + \beta r + 2\beta\mathbf{p}\right] u_0'(\mathbf{p}) \\
&\quad - \left[1 - \beta + \beta r + 2\beta\mathbf{p}\right] u_0'(\mathbf{p}) \\
\beta\left[-u_0(\mathbf{p}) + u_1\right] &= 0 \\
-u_0(\mathbf{p}) + u_1 &= 0 \\
u_1 &= u_0(\mathbf{p})
\end{aligned}$$

This is a contradiction of the set-up, because $u_1 > u_0(p)$, for all values of p in the choice set.

ii. Case ii. $p^P < p^M$

$$\begin{aligned}
\beta\left[u_0(p^M) - 2u_0(p^P) + u_1\right] &= \left[1 - \beta + \beta r + 2\beta p^M\right] u_0'(p^M) \\
&\quad - \left[1 - \beta + \beta r + 2\beta p^P\right] u_0'(p^P) \\
\beta\left[u_0(p^M) - 2u_0(p^P) + u_1\right] &= (1 - \beta + \beta r)\left[u_0'(p^M) - u_0'(p^P)\right] \\
&\quad + 2\beta\left(p^M - p^P\right)\right]\left[u_0'(p^M) - u_0'(p^P)\right]
\end{aligned}$$

$$\begin{aligned}
\left(u_0(p^M) - u_1\right) &\overset{?}{\leq} 2\left(u_0(p^P) - u_1\right) \\
u_0(p^M) + u_1 &\overset{?}{\leq} 2u_0(p^P)
\end{aligned}$$

If $p^P < p^M \Rightarrow u_0(p^P) > u_0(p^M)$. Furthermore, $u_1 > u_0(p^P)$, so it's a bit difficult to say what's going on with that sign yet (opposing effects).

On the other side,

$$p^P < p^M \Rightarrow \left[1 - \beta + \beta r + 2\beta p^M\right] > \left[1 - \beta + \beta r + 2\beta p^P\right].$$

So the coefficient on $u_0'(p^M)$ is larger than the coefficient on $u_0'(p^P)$.

a.) **No acceleration: $[\mathbf{u''} = \mathbf{0}]$**

Suppose $u'' = 0 \Rightarrow 0 > u_0'(p^M) = u_0'(p^P) \Rightarrow$

$$\begin{aligned}\beta\left[u_0(p^M) - 2u_0(p^P) + u_1\right] &= (1 - \beta + \beta r) \cdot 0 \\ &\quad + 2\beta\left(p^M u_0'(p^M) - p^P u_0'(p^P)\right) \\ \left[u_0(p^M) - 2u_0(p^P) + u_1\right] &= 2\left(p^M - p^P\right) u_0'(\mathbf{p})\end{aligned}$$

$p^M > p^P$ and $u_0' < 0 \Rightarrow$ RHS < 0.

This means RHS < 0, so LHS should be negative as well:

$$\begin{aligned}\beta\left[u_0(p^M) - 2u_0(p^P) + u_1\right] &< 0 \\ u_0(p^M) - 2u_0(p^P) + u_1 &< 0 \\ u_0(p^M) + u_1 &< 2u_0(p^P)\end{aligned}$$

Contradiction of assumptions.

b.) **Deceleration: $[\mathbf{u''} < \mathbf{0}]$**

Suppose $u'' < 0$ (i.e. concave, the function is slowing down). If $u'' < 0$, then $p^P < p^M \Rightarrow 0 > u_0'(p^P) > u_0'(p^M) \Rightarrow u_0'(p^M) - u_0'(p^P) < 0$. This also means that $p^M u_0'(p^M) - p^P u_0'(p^P) < 0$.

RHS is two negative terms added together so, RHS < 0. This means that LHS should also be < 0. Contradiction of assumptions.

c.) **Acceleration:** $[\mathbf{u}'' > \mathbf{0}]$

This is a contradiction of weak concavity.

$\square$

Proof of Lemma 3.4

Restatement of Lemma 3.4. *Under the assumptions listed, the following are the comparative statics of the problem:*

- *With respect to the discount factor, β, the non-cooperative action is increasing:*
$$\frac{\partial p^M}{\partial \beta} > 0,$$
as is the cooperative action:
$$\frac{\partial p^P}{\partial \beta} > 0.$$

- *With respect to the static good-state reward, u_1, the non-cooperative action is increasing:*
$$\frac{\partial p^M}{\partial u_1} > 0,$$
as is the cooperative action:
$$\frac{\partial p^P}{\partial u_1} > 0.$$

- *With respect to the transition from the good state to the bad state, r, the non-cooperative action is decreasing:*
$$\frac{\partial p^M}{\partial r} < 0,$$

while the cooperative action is constant:

$$\frac{\partial p^P}{\partial r} = 0.$$

Proof. This proof examines the comparative statics of the respective first order conditions using implicit function differentiation. First, I examine the comparative statics of the discount factor, β.

$$F(p^M) \equiv \beta \left(u_0(p^M) - u_1\right) - \left[1 - \beta + \beta r + 2\beta p^M\right] u_0'(p^M) = 0$$
$$G(p^P) \equiv 2\beta \left(u_0(p^P) - u_1\right) - \left[1 - \beta + \beta r + 2\beta p^P\right] u_0'(p^P) = 0$$

For the non-coordination problem:

$$\begin{aligned} \frac{\partial F(p^M)}{\partial \beta} &= \left(u_0(p^M) - u_1\right) + \beta u_0'(p^M)\frac{\partial p^M}{\partial \beta} - \left[-1 + r + 2p^m + 2\beta\frac{\partial p^M}{\partial \beta}\right] u_0'(p^M) \\ &\quad - \left[1 - \beta + \beta r + 2\beta p^M\right] u_0''(p^M)\frac{\partial p^M}{\partial \beta} = 0 \end{aligned}$$

$$\begin{aligned} &\left(u_0(p^M) - u_1\right) + \left(1 - r - 2p^M\right) u_0'(p^M) \\ &\qquad = \left[1 - \beta + \beta r + 2\beta p^M\right] u_0''(p^M)\frac{\partial p^M}{\partial \beta} + 2\beta\frac{\partial p^M}{\partial \beta}u_0'(p^M) - \beta u_0'(p^M)\frac{\partial p^M}{\partial \beta} \\ &\left(u_0(p^M) - u_1\right) + \left(1 - r - 2p^M\right) u_0'(p^M) \\ &\qquad = \left[\left(1 - \beta + \beta r + 2\beta p^M\right) u_0''(p^M) + 2\beta u_0'(p^M) - \beta u_0'(p^M)\right] \frac{\partial p^M}{\partial \beta} \\ &\left(u_0(p^M) - u_1\right) + \left(1 - r - 2p^M\right) u_0'(p^M) \\ &\qquad = \left[\left(1 - \beta + \beta r + 2\beta p^M\right) u_0''(p^M) + \beta u_0'(p^M)\right] \frac{\partial p^M}{\partial \beta} \\ &\frac{\partial p^M}{\partial \beta} = \frac{\left(u_0(p^M) - u_1\right) + \left(1 - r - 2p^M\right) u_0'(p^M)}{\left[\left(1 - \beta + \beta r + 2\beta p^M\right) u_0''(p^M) + \beta u_0'(p^M)\right]} \end{aligned}$$

From the earlier assumptions, the denominator is not equal to zero. Furthermore it is negative.

$$\frac{\partial p^M}{\partial \beta} \overset{?}{<} 0$$
$$\frac{\left(u_0(p^M) - u_1\right) + \left(1 - r - 2p^M\right) u_0'(p^M)}{\left[(1 - \beta + \beta r + 2\beta p^M)\, u_0''(p^M) + \beta u_0'(p^M)\right]} \overset{?}{<} 0$$

To find the sign of this expression, need to know when the numerator is negative:

$$\begin{aligned}
\left(u_0(p^M) - u_1\right) + \left(1 - r - 2p^M\right) u_0'(p^M) &\overset{?}{<} 0 \\
-\left(u_1 - u_0(p^M)\right) + \left(1 - r - 2p^M\right) u_0'(p^M) &\overset{?}{<} 0 \\
\left(1 - r - 2p^M\right) u_0'(p^M) &\overset{?}{<} \left(u_1 - u_0(p^M)\right)
\end{aligned}$$

The numerator is negative when $1-r-2p^M$ is positive or if the following holds:

$$0 > \left(1 - r - 2p^M\right) > \frac{\left(u_1 - u_0(p^M)\right)}{u_0'(p^M)} \tag{D.1}$$

For the coordination problem:

$$\begin{aligned}
\frac{\partial G(p^P)}{\partial \beta} =& 2\left(u_0(p^P) - u_1\right) + 2\beta u_0'(p^P)\frac{\partial p^P}{\partial \beta} - \left[-1 + r + 2p^P + 2\beta\frac{\partial p^P}{\partial \beta}\right] u_0'(p^P) \\
&- \left[1 - \beta + \beta r + 2\beta p^P\right] u_0''(p^P)\frac{\partial p^P}{\partial \beta} = 0
\end{aligned}$$

$$
\begin{aligned}
&2\left(u_0(p^P)-u_1\right)+\left(1-r-2p^P\right)u_0'(p^P)\\
&\qquad=\left[1-\beta+\beta r+2\beta p^P\right]u_0''(p^P)\frac{\partial p^P}{\partial \beta}-2\beta u_0'(p^P)\frac{\partial p^P}{\partial \beta}+2\beta\frac{\partial p^P}{\partial \beta}u_0'(p^P)\\
&2\left(u_0(p^P)-u_1\right)+\left(1-r-2p^P\right)u_0'(p^P)=\left[1-\beta+\beta r+2\beta p^P\right]u_0''(p^P)\frac{\partial p^P}{\partial \beta}\\
&\qquad\frac{\partial p^P}{\partial \beta}=\frac{2\left(u_0(p^P)-u_1\right)+\left(1-r-2p^P\right)u_0'(p^P)}{\left[1-\beta+\beta r+2\beta p^P\right]u_0''(p^P)}
\end{aligned}
$$

The denominator is negative. Checking when the numerator is negative:

$$
\begin{aligned}
2\left(u_0(p^P)-u_1\right)+\left(1-r-2p^P\right)u_0'(p^P)&\overset{?}{<}0\\
-2\left(u_1-u_0(p^P)\right)+\left(1-r-2p^P\right)u_0'(p^P)&\overset{?}{<}0\\
\left(1-r-2p^P\right)u_0'(p^P)&\overset{?}{<}2\left(u_1-u_0(p^P)\right)
\end{aligned}
$$

If Equation (D.1) holds, then this holds more easily, since $p^P > p^M$.

Now I examine the comparative statics of the good-state reward, u_1.

$$
\begin{aligned}
F(p^M)&\equiv\beta\left(u_0(p^M)-u_1\right)-\left[1-\beta+\beta r+2\beta p^M\right]u_0'(p^M)=0\\
G(p^P)&\equiv2\beta\left(u_0(p^P)-u_1\right)-\left[1-\beta+\beta r+2\beta p^P\right]u_0'(p^P)=0
\end{aligned}
$$

For the non-coordination problem:

$$
\begin{aligned}
\frac{\partial F(p^M)}{\partial u_1}=&\beta\left(u_0'(p^M)\frac{\partial p^M}{\partial u_1}-1\right)-2\beta\frac{\partial p^M}{\partial u_1}u_0'(p^M)\\
&-\left[1-\beta+\beta r+2\beta p^M\right]u_0''(p^M)\frac{\partial p^M}{\partial u_1}=0
\end{aligned}
$$

$$
\begin{aligned}
\left[\beta u_0'(p^M)-2\beta u_0'(p^M)-\left(1-\beta+\beta r+2\beta p^M\right)u_0''(p^M)\right]\frac{\partial p^M}{\partial u_1}&=\beta\\
\left[-\beta u_0'(p^M)-\left(1-\beta+\beta r+2\beta p^M\right)u_0''(p^M)\right]\frac{\partial p^M}{\partial u_1}&=\beta
\end{aligned}
$$

$$\frac{\partial p^M}{\partial u_1} = \frac{-\beta}{\left[\beta u_0'(p^M) + (1-\beta+\beta r + 2\beta p^M)\, u_0''(p^M)\right]}$$

Given the previous assumptions, the denominator is negative, rendering the whole thing positive.

For the coordination problem:

$$\begin{aligned}\frac{\partial G(p^P)}{\partial u_1} =& 2\beta\left(u_0'(p^P)\frac{\partial p^P}{\partial u_1} - 1\right) - 2\beta\frac{\partial p^P}{\partial u_1}u_0'(p^P) \\ & - \left[1-\beta+\beta r + 2\beta p^P\right] u_0''(p^P)\frac{\partial p^P}{\partial u_1} = 0\end{aligned}$$

$$\begin{aligned}2\beta u_0'(p^P)\frac{\partial p^P}{\partial u_1} - 2\beta\frac{\partial p^P}{\partial u_1}u_0'(p^P) - \left[1-\beta+\beta r + 2\beta p^P\right] u_0''(p^P)\frac{\partial p^P}{\partial u_1} &= 2\beta \\ - \left[1-\beta+\beta r + 2\beta p^P\right] u_0''(p^P)\frac{\partial p^P}{\partial u_1} &= 2\beta\end{aligned}$$

$$\frac{\partial p^P}{\partial u_1} = \frac{-2\beta}{\left[1-\beta+\beta r + 2\beta p^P\right] u_0''(p^P)}$$

This is positive. In comparing the two:

$$\begin{aligned}\frac{\partial p^M}{\partial u_1} &= \frac{-\beta}{\left[\beta u_0'(p^M) + (1-\beta+\beta r + 2\beta p^M)\, u_0''(p^M)\right]} \\ \frac{\partial p^P}{\partial u_1} &= \frac{-2\beta}{\left[1-\beta+\beta r + 2\beta p^P\right] u_0''(p^P)}\end{aligned}$$

The denominator is smaller, the numerator is larger, so the affect of u_1 on p^P is much greater than that on p^M.

Finally, I examine the comparative statics of the transition probability from the good state to the bad state, r.

$$F(p^M) \equiv \beta\left(u_0(p^M) - u_1\right) - \left[1 - \beta + \beta r + 2\beta p^M\right] u_0'(p^M) = 0$$
$$G(p^P) \equiv 2\beta\left(u_0(p^P) - u_1\right) - \left[1 - \beta + \beta r + 2\beta p^P\right] u_0'(p^P) = 0$$

For the non-coordination problem:

$$\begin{aligned}\frac{\partial F(p^M)}{\partial r} =& \beta u_0'(p^M)\frac{\partial p^M}{\partial r} - \beta u_0'(p^M) \\ & - \left[1 - \beta + \beta r + 2\beta p^M\right] u_0''(p^M)\frac{\partial p^M}{\partial r} = 0\end{aligned}$$

$$\frac{\partial p^M}{\partial r} = \frac{\beta u_0'(p^M)}{\left[\beta u_0'(p^M) - (1 - \beta + \beta r + 2\beta p^M)\, u_0''(p^M)\right]}$$

The numerator is negative; the denominator is positive if:

$$\left[\beta u_0'(p^M) - \left(1 - \beta + \beta r + 2\beta p^M\right) u_0''(p^M)\right] \overset{?}{>} 0$$
$$0 > \beta u_0'(p^M) \overset{?}{>} \left(1 - \beta + \beta r + 2\beta p^M\right) u_0''(p^M)$$

For the coordination problem:

$$\begin{aligned}\frac{\partial G(p^P)}{\partial r} =& 2\beta u_0'(p^P)\frac{\partial p^P}{\partial r} - 2\beta\frac{\partial p^P}{\partial r}u_0'(p^P) \\ & - \left[1 - \beta + \beta r + 2\beta p^P\right] u_0''(p^P)\frac{\partial p^P}{\partial r} = 0\end{aligned}$$

$$\begin{aligned}- \left[1 - \beta + \beta r + 2\beta p^P\right] u_0''(p^P)\frac{\partial p^P}{\partial r} &= 0 \\ \frac{\partial p^P}{\partial r} &= 0\end{aligned}$$

□

Bibliography

[1] Fish Stock Assessment 101. Technical report, NOAA Fisheries.
Cited on page(s): 97

[2] Tuvalu's views on the possible security implications of climate change to be included in the report of the UN secretary general to the UN general assembly 64th session, November 2009.
Cited on page(s): 13

[3] Emission database for global atmospheric research (EDGAR), 2011. Release version 4.2.
Cited on page(s): 12

[4] Europe.eu: Money and the EU, January 2013.
Cited on page(s): 2

[5] Europe.eu: The history of the European Union, January 2013.
Cited on page(s): 2

[6] FishstatJ – Software for global capture production 1950-2012, 2014.
Cited on page(s): 14

[7] Fishwatch.gov: Pacific Albacore tuna, March 2014.
Cited on page(s): 14

[8] Robert M. Anderson and William R. Zame. Genericity with infinitely many parameters. *Advances in Theoretical Economics*, 1(1):1–64, 2001.
Cited on page(s): 33

[9] Kyle Bagwell. Self-enforcing trade agreements and private information. Technical report, National Bureau of Economic Research, 2009.
Cited on page(s): 10

[10] Scott Barrett. Self-enforcing international environmental agreements. *Oxford Economic Papers*, 46:878–894, 1994.
Cited on page(s): 5

[11] Johann D. Bell, Chris Reid, Michael J. Batty, Patrick Lehodey, Len Rodwell, Alistair J. Hobday, Johanna E. Johnson, and Andreas Demmke. Effects of climate change on oceanic fisheries in the tropical Pacific: Implications for economic development and food security. *Climatic Change*, 119(1):199–212, 2013.
Cited on page(s): 15

[12] Alain L. Bernard and Marc Vielle. Measuring the welfare cost of climate change policies: A comparative assessment based on the computable general equilibrium model GEMINI-E3. *Environmental Modeling & Assessment*, 8(3):199–217, 2003.
Cited on page(s): 63

[13] 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.

Cited on page(s): 7, 37

[14] Trond Bjørndal. Production economics and optimal stock size in a North Atlantic fishery. *The Scandinavian Journal of Economics*, 89(2):145, June 1987.

Cited on page(s): 65

[15] 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.

Cited on page(s): 5

[16] Marek Błaś, Mieczysław Sobik, Joanna Potocka, and Ryszard Twarowski. Alkalizacja hydrometeorów w Sudetach Zachodnich. In *Sborník z Mezinárodní Vědecké Konference*, volume 41 of *Opera Corcontica*, pages 74–85, Szklarska Poręba, November 2003.

Cited on page(s): 17

[17] Thomas Brunel and Jean Boucher. Long-term trends in fish recruitment in the north-east Atlantic related to climate change. *Fisheries Oceanography*, 16(4):336–349, 2007.

Cited on page(s): 97

[18] Reniel B. Cabral, Rollan C. Geronimo, May T. Lim, and Porfirio M. Aliño. Effect of variable fishing strategy on fisheries under changing effort

and pressure: An agent-based model application. *Ecological Modelling*, 221(2):362–369, January 2010.
Cited on page(s): 97

[19] Kenneth I. Carlaw and Richard G. Lipsey. Externalities, technological complementarities and sustained economic growth. *Research Policy*, 31(8):1305–1315, 2002.
Cited on page(s): 63

[20] Carlo Carraro, Carmen Marchiori, and Sonia Oreffice. Endogenous minimum participation in international environmental treaties. *Environmental and Resource Economics*, 42(3):411–425, 2009.
Cited on page(s): 6, 28

[21] Arthur Charpentier and Benoît Le Maux. Natural catastrophe insurance: How should the government intervene? *Journal of Public Economics*, 115:1–17, July 2014.
Cited on page(s): 68

[22] Ward Cheney. Calculus in Banach spaces. In *Analysis for Applied Mathematics*, volume 208 of *Graduate Texts in Mathematics*, pages 115–169. Springer-Verlag, 2001.
Cited on page(s): 156

[23] Colin W. Clark. *The worldwide crisis in fisheries: Economic models and human behavior*. Cambridge University Press, 2006.
Cited on page(s): 95

[24] Simone Clemhout and H. Y. Wan. Dynamic common property resources and environmental problems. *Journal of Optimization Theory and Applications*, 46(4):471–481, 1985.

Cited on page(s): 64

[25] James E. Cloern and Frederic H. Nichols. A von Bertalanffy growth model with a seasonally varying coefficient. *Journal of the Fisheries Research Board of Canada*, 35:1479–1482, 1978.

Cited on page(s): 97

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

Cited on page(s): 139

[27] Manjira Datta and Leonard J. Mirman. Externalities, market power, and resource extraction. *Journal of Environmental Economics and Management*, 37(3):233–255, 1999.

Cited on page(s): 66

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

Cited on page(s): 7

[29] Étienne Billette de Villemeur and Justin Leroux. Looking ahead versus

looking back: Revisiting the carbon tax!, 2011.

Cited on page(s): 41

[30] Einar Eythórsson. Theory and practice of ITQs in Iceland. Privatization of common fishing rights. *Marine Policy*, 20(3):269–281, 1996.

Cited on page(s): 98

[31] Y. Hossein Farzin. Optimal pricing of environmental and natural resource use with stock externalities. *Journal of Public Economics*, 62(1):31–57, 1996.

Cited on page(s): 67

[32] Ronald D. Fischer and Leonard J. Mirman. Strategic dynamic interaction: Fish wars. *Journal of Economic Dynamics and Control*, 16(2):267–287, 1992.

Cited on page(s): 63, 65

[33] International Union for the Conservation of Nature + Natural Resources IUCN. Environmental Treaty Status Data Set, 2012 Release, 2012.

Cited on page(s): 95

[34] Rainer Froese. Keep it simple: Three indicators to deal with overfishing. *Fish and Fisheries*, 5(1):86–91, 2004.

Cited on page(s): 98

[35] Emily Galloway and Erik Paul Johnson. Teaching an old dog new tricks: Firm learning from environmental regulation. 2014.
Cited on page(s): 63

[36] David Griffiths and Chris Harrod. Natural mortality, growth parameters, and environmental temperature in fishes revisited. *Canadian Journal of Fisheries and Aquatic Sciences*, 64:249–255, 2007.
Cited on page(s): 97

[37] Alexander Hamilton. Federalist paper no. 1. 1787.
Cited on page(s): 2

[38] Ling Huang and Martin D. Smith. The dynamic efficiency costs of common-pool resource exploitation. *American Economic Review*, 104(12):4071–4103, December 2014.
Cited on page(s): 87, 96, 97

[39] Daniel D. Huppert. An overview of fishing rights. *Reviews in Fish Biology and Fisheries*, 15(3):201–215, August 2005.
Cited on page(s): 97, 98

[40] Apisai Ielemia. Statement by Tuvalu, September 2009.
Cited on page(s): 13

[41] James Joseph, W.L. Klawe, Pat Murphy, Inter-American Tropical Tuna Commission, et al. *Tuna and billfish: Fish without a country.* Inter-

American Tropical Tuna Commission, 1980.
Cited on page(s): 95

[42] Athanasios Kampas. On the Allocation of Possible EU Total Allowable Catches (TAC) for the Mediterranean Swordfish: An Envy-Free Criterion and Equitable Procedure. *Journal of Agricultural Economics*, pages 1–22, February 2014.
Cited on page(s): 98

[43] 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.
Cited on page(s): 7

[44] Roger Lagunoff and Rodrigo Harrison. Dynamic mechanism design for a global commons. *Available at SSRN*, 2013.
Cited on page(s): 9

[45] Rense Lange, Sean McDade, and Terence A. Oliva. Technological choice and network externalities: A catastrophe model analysis of firm software adoption for competing operating systems. *Structural Change and Economic Dynamics*, 12(1):29–57, 2001.
Cited on page(s): 69

[46] David Levhari and Leonard J. Mirman. The great fish war: An example using a dynamic Cournot-Nash solution. *The Bell Journal of Economics*,

11(1):322, 1980.

Cited on page(s): 97

[47] O. Yale Lewis. Treaty fishing rights: A habitat right as part of the trinity of rights implied by the fishing clause of the Stevens Treaties. *American Indian Law Review*, 27(1):281, 2002.

Cited on page(s): 97

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

Cited on page(s): 10, 38

[49] Tsujimura Motoh. Optimal natural resources management under uncertainty with catastrophic risk. *Energy Economics*, 26(3):487–499, May 2004.

Cited on page(s): 68

[50] Ransom A. Myers and Ken Drinkwater. The influence of Gulf Stream warm core rings on recruitment of fish in the northwest Atlantic. *Journal of Marine Research*, 47(3):635–656, 1989.

Cited on page(s): 97

[51] Jos G. J. Olivier, Greet Janssens-Maenhout, Marilena Muntean, and Jeroen A. H. W. Peters. Trends in global CO_2 emissions: 2013 report. Technical report, The Hague: PBL Netherlands Environmental Assessment Agency; Ispra: Joint Research Centre, 2013.

Cited on page(s): 12

[52] Waidi Oyebanjo Abdul, Isaac Tunde Omoniyi, Yemi Akegbejo-Samsons, Adeolu A. Akinyemi, Adrew O. Agbon, and Festus Idowu Adeosun. Management indicators and growth performance index of *Tilapia zillii* in a tropical coastal estuary. *Journal of Agricultural Science*, 4(11), October 2012.

Cited on page(s): 65

[53] Daniel Pauly and Maria-Lourdes Palomares. Fishing down marine food web: It is far more pervasive than we thought. *Bulletin of Marine Science*, 76(2):197–212, 2005.

Cited on page(s): 95

[54] Peter H. Pearse and Carl J. Walters. Harvesting regulation under quota management systems for ocean fisheries. *Marine Policy*, 16(3):167–182, May 1992.

Cited on page(s): 98

[55] Katharine L. Ricke, Juan B. Moreno-Cruz, and Ken Caldeira. Strategic incentives for climate geoengineering coalitions to exclude broad participation. *Environmental Research Letters*, 8(1), 2013.

Cited on page(s): 11

[56] William E. Ricker. Stock and recruitment. *Journal of the Fisheries Board of Canada*, 11(5):559–623, 1954.

Cited on page(s): 65, 97

[57] Hilary Sigman. International spillovers and water quality in rivers: Do countries free ride? *The American Economic Review*, 92(4):1152–1159, 2002.

Cited on page(s): 16

[58] James E. Smith and Kevin F. McCardle. Structural properties of stochastic dynamic programs. *Operations Research*, 50(5):796–809, October 2002.

Cited on page(s): 99

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

Cited on page(s): 7, 9, 38

[60] Martin L. Weitzman. On modeling and interpreting the economics of catastrophic climate change. *The Review of Economics and Statistics*, 91(1), 2009.

Cited on page(s): 68

[61] World Bank. *World Development Indicators 2013*. The World Bank, 2013.

Cited on page(s): 12

Vita

Anna Alexandra Klis was born in the mountains of Colorado in the year 1989. Her parents and she lived in a number of states during her childhood, though her primary schooling was in Toledo, Ohio. After completing her secondary education in The Woodlands, Texas, Anna attended Georgetown University's Edmund A. Walsh School of Foreign Service. In three years, she earned her Bachelor of Science in Foreign Service, with honors in the International Economics major, while participating in student theatre with the Mask & Bauble Dramatic Society, volunteering with Learning Enterprises as a summer teacher and Program Director in Poland, and serving on the steering committee of the Carroll Round as Editor-in-Chief of the Proceedings. Upon graduation from Georgetown, Anna entered graduate studies in the Department of Economics at the University of Texas at Austin in August, 2010, to continue to advance her studies in the economics of international relations.

Email address: aklis@utexas.edu

This dissertation was typeset with LaTeX[†] by the author.

[†] LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.