INTERNALIZING EXTERNALITIES: ROLES OF NETWORKS, CLUBS AND POLICY COMMITMENTS

A Dissertation
Presented to
The Academic Faculty

By

Jingwen Qu

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Economics

Georgia Institute of Technology

May 2020

INTERNALIZING EXTERNALITIES: ROLES OF NETWORKS, CLUBS AND POLICY COMMITMENTS

Approved by:

Dr. Tibor Besedes
School of Economics
Georgia Institute of Technology

Dr. Justin Burkett
School of Economics
Georgia Institute of Technology

Dr. Matthew Oliver
School of Economics
Georgia Institute of Technology

Dr. Emilson Silva
Department of Marketing, Business Economics and Law
University of Alberta

Dr. Sudipta Sarangi
Department of Economics
Virginia Tech

Date Approved: March 23, 2020

To my parents

ACKNOWLEDGEMENTS

I am grateful to Tibor Besedes for his supervision and support. I am also grateful to Tibor Besedes, Justin Burkett, Matthew Oliver, Emilson Silva and Sudipta Sarangi for serving on my committee and giving invaluable comments and suggestions.

I would like to express my gratitude to Emilson Silva for the many enlightening discussions and support over the years. I also thank Mikhail Klimenko, Levent Kutlu and Derek Tittle for introducing me to various topics. And I have been fortunate to meet other great scholars during conferences, seminars, lectures, etc. Moreover, I am grateful to everyone that has helped and encouraged me.

Lastly, I would like to express my gratitude to my parents, for everything.

TABLE OF CONTENTS

Acknowledgments v

List of Tables x

List of Figures xi

Chapter 1: Introduction 1

Chapter 2: Voluntary Provision of Public Goods in Endogenous Networks: A Perspective from Social Incentives 3

2.1 Introduction 3

2.2 Relation to the Literature 5

2.3 Model Setup 7

2.4 Shape of Equilibrium Profiles 10

2.4.1 One-way Flows 11

2.4.2 Two-way Flows 12

2.4.3 The Law of the Few 14

2.5 Effects of Social Incentives 15

2.5.1 Do Social Incentives Always Induce More Volunteers and Who Are the Volunteers? 16

2.5.2 Do Social Incentives Make Everyone Better Off? 19

2.6 Extensions 20
2.6.1 Heterogeneity in Consumption Benefits and Production Costs of the Public Good as well as Linking Costs 20
2.6.2 Imperfect Substitutability Between One's Own and Others' Volunteering Efforts 22
2.7 Conclusion 23

Chapter 3: Multiple Climate Clubs 24
3.1 Introduction 24
3.1.1 Related Literature 27
3.2 Basic Model 28
3.3 Equilibrium Analysis 30
3.3.1 Stage Three 30
3.3.2 Stage Two 31
3.3.3 Stage One 34
3.4 Extension 36
3.5 Summary 40
3.5.1 Equilibrium Stability and Efficiency 40
3.5.2 Effects of the Formation of Multiple Clubs 43
3.5.3 Effects of FTAs 44

Chapter 4: Strategic Effects of Future Environmental Policy Commitments: Climate Change, Solar Radiation Management and Correlated Air Pollutants 46
4.1 Introduction 46
4.2 Modeling Strategies and Brief Discussion of Main Results 49

4.3 General Equilibrium Framework 51
4.4 Effects of Environmental Policy Commitments 58
4.4.1 Game I: Simultaneous Choices of Carbon Quotas and SRM Levels . 61
4.4.2 Game II: Simultaneous Coordinated Choices 62
4.4.3 Game III: Simultaneous Commitments on Carbon Permit Policies . 63
4.4.4 Game IV: Simultaneous Commitments on SRM Policies 64
4.4.5 Game V: National Leadership in Carbon Permit Policy 65
4.4.6 Game VI: National Leadership in SRM Policy 66
4.5 First-mover Advantage 66
4.6 Comparisons across Equilibria 68
4.6.1 Symmetric Economy 68
4.6.2 Robustness: Asymmetric Drought and Acid Rain Damages 73
4.7 Conclusion 75

Appendix A: Appendix for Chapter 2 79
A.1 Proofs 79
A.1.1 Proof of Proposition 2.1 79
A.1.2 Proof of Proposition 2.2 80
A.1.3 Proof of Corollary 2.1 82
A.1.4 Proof of Lemma 2.1 83
A.1.5 Proof of Proposition 2.3 83
A.1.6 Proof of Proposition 2.4 83
A.1.7 Proof of Remark 2.1 86

A.1.8 Proof of Proposition 2.5 87

A.1.9 Proof of Proposition 2.6 88

A.1.10 Proof of Proposition 2.7 90

A.1.11 Proof of Proposition 2.8 90

A.1.12 Proofs of Propositions 2.9 and 2.10 91

A.2 Dynamic Setup for Section 2.5.1 91

Appendix B: Appendix for Chapter 3 94

B.1 Proofs 94

B.1.1 Proof of Result 3.3 94

B.1.2 Proof of Result 3.4 94

B.1.3 Proof of Result 3.8 95

B.2 Tables 97

References 121

LIST OF TABLES

4.1 Environmental policy games. 59

4.2 Main results of baseline comparisons across equilibria. 69

4.3 Rankings with asymmetric drought and acid rain damages. 77

B.1 Rankings of national welfare in the absence of FTAs ($n = 3$). 98

B.2 Rankings of national welfare in the presence of FTAs ($n = 3$). 99

B.3 Equilibria in the absence of FTAs ($n = 3$). 100

B.4 Equilibria in the presence of FTAs ($n = 3$). 101

B.5 Rankings of national welfare in the absence of FTAs ($n = 4$). 102

B.6 Rankings of national welfare in the presence of FTAs ($n = 4$). 104

B.7 Equilibria in the absence of FTAs ($n = 4$). 107

B.8 Equilibria in the presence of FTAs ($n = 4$). 109

B.9 Rankings of global welfare across trade regimes ($n = 3$). 111

B.10 Rankings of global welfare across trade regimes ($n = 4$). 113

LIST OF FIGURES

2.1 Examples of strict Nash equilibria for one-way flows. 11

2.2 Examples of strict Nash equilibria for two-way flows. 13

2.3 Examples of stochastically stable equilibria for one-way and two-way flows. 18

3.1 Possible club structures resulted from the first stage ($n = 3$). 31

3.2 Possible club structures resulted from the first stage ($n = 4$). 37

SUMMARY

Externalities arising from actions of one player in the economy and directly affecting the well-being of another are ubiquitous. In these situations, market equilibria often fail to be efficient. This dissertation explores roles of networks, clubs and policy commitments in the internalization of externalities and thus in the generation of efficient outcomes. The first essay examines how network-based social incentives would affect the provision of public goods in endogenous networks. The second essay analyzes effects of the formation of multiple climate clubs and free trade agreements among club members on the stability and efficiency of international environmental agreements. The third essay studies effects of policy commitments to the provision of a new global-warming-relieving technology called solar radiation management relative to effects of policy commitments to carbon mitigation.

CHAPTER 1

INTRODUCTION

It is a ubiquitous phenomenon that actions of one player in the economy may produce externalities directly affecting the well-being of another. In these situations, market equilibria usually fail to yield efficient outcomes. Various suggestions have been proposed for alternative ways that would lead to efficiency, including taxes, subsidies, quotas, etc. This dissertation builds upon these traditional solutions and explores the roles of networks, clubs and policy commitments in the internalization of externalities and thus the generation of efficiency.

The first essay examines the role of social incentives in public goods games in endogenous networks. Examples of such network-based incentives include feelings of pleasure associated with helping others, peer recognition, recipient-based material rewards, etc. We find that large social benefits will always induce more volunteers only when the network flow of benefits is two-way. In these cases, volunteers could initiate links with free riders. For both one-way and two-way flows, volunteers are not necessarily those who enjoy the largest social benefits from each recipient. We also characterize the condition that ensures efficient provision of public goods in a stochastically stable equilibrium of either flow type: the social benefits from each recipient are large relative to the linking cost for sufficiently many individuals.

The second essay analyzes effects of the formation of multiple climate clubs and free trade agreements among club members on the stability and efficiency of international environmental agreements. The major takeaways are that the formation of multiple climate clubs improves equilibrium stability when emissions coefficients are not large, and the inclusion of free trade in club treaties enlarges the stable and efficient region of the equilibrium. Meanwhile, stable club structures could involve more than three countries as participants of the clubs. In particular, the grand coalition would be the only stable and efficient structure when club members free trade with each other and emissions coefficients are sufficiently large. Moreover, we find that stable equilibrium structures based on coalition-proofness and group stability would be different if multiple clubs are formed and free trade is included.

The third essay studies effects of environmental policy commitments in a futuristic world in which solar radiation management (SRM) can be utilized to reduce climate change damages. Carbon and sulfur dioxide emissions (correlated pollutants) can be reduced through tradable permits. We show that if nations simultaneously commit to carbon permit policies, national SRM levels rise with carbon quotas. Alternatively, if they simultaneously commit to SRM policies, the global temperature falls with each unit increase in the global SRM level. A nation always wishes to be a leader in policymaking, but prefers carbon to SRM policymaking. The globe prefers SRM policy commitments.

CHAPTER 2

VOLUNTARY PROVISION OF PUBLIC GOODS IN ENDOGENOUS NETWORKS: A PERSPECTIVE FROM SOCIAL INCENTIVES

2.1 Introduction

Individuals commonly get access to local public goods through personal acquisition and social networking. For instance, consumers learn about product features by making own searches and talking with market mavens; users of open-source software or contents may make their own contribution to the projects and benefit from others' efforts through communication; and farmers usually get familiar with new technologies through own experimentation and discussions with each other. Both personal acquisition and social networking consume time and resources. Empirical and experimental studies have acknowledged *social incentives* as important motivations for personal provision of public goods in endogenously formed social networks.[1] Examples of such network-based incentives include: feelings of pleasure associated with helping others, peer recognition, recipient-based material rewards, etc. [53, 40, 55, 11].

Several questions remain regarding the role of social incentives in the provision of public goods in endogenous networks. Do such incentives always induce more volunteers? For those who choose to free ride, is it because they derive less payoffs from the incentives than volunteers do, *ceteris paribus*? And what are the welfare effects?

This paper makes a first attempt to address these questions by integrating the social incentives into [31]'s model (henceforth $G\&G$) of local public goods games in endogenous networks. Our analysis also yields insights for designing environments to induce efficient

[1]In our first example, the underlying networks are endogenous as consumers could establish social contacts with market mavens to learn about product features; conversely, market mavens could also form links with consumers to share their knowledge about the products. The social networks are endogenous in a similar sense in the other two examples.

decisions on public goods provision and network formation.

More specifically, players choose whether or not to provide a public good and with whom to form social links. They do so by taking into account social benefits of volunteering, which depend on the numbers of recipients of the public goods they provide. In addition, we consider both *one-way* and *two-way flow of benefits* in networks.[2] In the former case, only the sponsor of a link could access the public good provided by the other person involved[3], while in the latter case, the transmission of the public good could go in either direction no matter who incurs the linking cost[4].

The main findings are as follows. First, when the flow of benefits is one-way, every strict Nash equilibrium has a *core-periphery multi network* structure. Links between volunteers and free riders all go from the latter to the former. Every player would be better off in an equilibrium with more volunteers. However, the presence of social incentives would not always bring more Nash volunteers, no matter how large the benefits are. This is due to the fact that for one-way flows, a volunteer's social benefits only depend on the number of incoming links from others. It is then possible for coordination failure to occur.

Second, when the flow of benefits is two-way, a strict Nash equilibrium has a *nested split network* structure, which is either a *core-periphery simple network* or a *multipartite simple network*. A volunteer's social benefits now depend on the total number of incoming and outgoing links. Large enough social incentives could always induce more Nash volunteers by incentivizing them to initiate links with others, which in turn increases their social benefits. Having more volunteers would make everyone better off as long as there are not too many of them or the linking cost is sufficiently low relative to the social benefits from each recipient.

Third, for either type of network flow, Nash volunteers are not necessarily those who

[2][6], [30], [32] and [31], among others, have studied one-way and two-way flow of benefits in networks.

[3]For example, one GitHub user follows another so as to get notifications about the latter's latest public activities such as uploading new codes.

[4]For example, consumers could learn about features of new products from the phone calls made by market mavens.

enjoy the largest social benefits from each recipient. This can be seen from the fact that the social benefits of volunteering depend on the number of incoming (and outgoing) links, and link formation is costly. Therefore, it is always possible that some players choose to free ride just due to too few incoming links.

Fourth, in a stochastically stable equilibrium of either type of flow, social incentives could always bring more volunteers and make everyone better off. This happens when the social benefits from each recipient of the public goods are sufficiently large relative to the linking cost for sufficiently many individuals.

Finally, in large societies, every strict Nash equilibrium of either type of network flow has a *core-periphery structure* and a very small proportion of the population provide the public goods. That is, every equilibrium exhibits the empirically robust *law of the few*, subsumed by $G\&G$. The result obtains due to the joint forces of several factors: links are endogenous and costly; the public goods provided by oneself and by neighbors are substitutable; different individuals generate the same value of social benefits, *ceteris paribus*; and marginal returns to both consumption and social benefits are decreasing.

The organization of the rest of the chapter is as follows. Section 3.2 discusses our contribution to the literature. Section 3 describes model setup. Section 3.4 characterizes (strict) Nash equilibria. Section 3.5 analyzes effects of social incentives on the equilibria. Section 3.6 presents two extensions of the basic model. Section 3.7 concludes.

2.2 Relation to the Literature

The paper contributes to the literature by introducing social incentives into a model of local public goods game in endogenous networks. Such incentives are widely acknowledged in models of network-free public goods games and other prosocial activities (see, among others, [19], [2], [3] and [10]). To the best of our knowledge, current models studying local public good games in endogenous/fixed networks (see, among others, [14], [31], [1] and [37]) have not considered such incentives as motivations for public goods provision.

One exception is [39], who makes a similar attempt by considering status rents generated from attractive network positions in public goods games on one-way flow networks. Our paper differs from theirs in two aspects: i) the social incentives in our model depend on individual decisions on both volunteering and networking; ii) we study both one-way and two-way flow of benefits and examine the differences and similarities in the way social incentives work to affect the equilibria in these two types of networks. Other related but different works, for example, [13], [23], [35] and [46], consider altruism, reciprocity, social status or warm glow on fixed networks.

This exercise helps to enhance our understanding of the effects of social incentives on public goods provision. In network-free settings, the consideration of such incentives always leads to more provision.[5] This, however, is not necessarily the case when the public goods and the incentives are embedded in endogenous networks. If everyone plays best response in a one-shot game, social incentives will always bring more volunteers only if the network flow of benefits is two-way. In these cases, sufficiently large incentives would induce volunteers to actively form links with others, which in turn increase their social benefits and thus stabilizing their volunteering behavior. Alternatively, if individuals respond periodically and myopically and the dynamics of the play follows logit quantal response rules, then the occasional deviations due to say mistakes or experimentation would allow players to coordinate on the equilibrium with more volunteers for both types of flows.

We note that the active linking behavior of volunteers with two-way flows have not been reported in previous studies without the consideration of social incentives. In these setups, equilibrium links between volunteers and free riders all go from the latter to the former (see, among others, [31] and [37]). It is not uncommon, however, to observe such active linking behavior in real life. For example, market mavens often initiate conversations with other consumers to inform them about product features or places to shop [28]. And when a group of people are planning for a road trip, warm-hearted individuals usually actively

[5] Our comparison of the network-free and the current setup focuses on the overall effects of social incentives rather than how different types of the incentives interact with each other.

keep others updated about their findings on optimal routes, etc. The incorporation of social incentives therefore speaks to these additional phenomena.

Moreover, our welfare analysis yields insights for designing environments that could induce efficient voluntary provision of public goods in endogenous networks. If the network flow of benefits is one-way, then the larger the social incentives from each recipient for sufficiently many people, the better off everyone will be in a stochastically stable equilibrium. If the flow of benefits is two-way, keeping the cost of linking low enough relative to social benefits per recipient would make everyone better off with larger incentives in every strict Nash equilibrium.

2.3 Model Setup

We study a public good game in an endogenous network, with the consideration of social incentives. Let $N = \{1, 2, ..., n\}$ with $n \geq 3$ be a set of players. Player $i \in N$ simultaneously determines her volunteering strategy $x_i \in \{0, 1\}$ and linking strategies $g_{ij} \in \{0, 1\}, \forall j \in N \backslash \{i\}$. We call player i a *volunteer* if she chooses to provide the public good ($i.e.$, $x_i = 1$) and a *free rider* otherwise. For player $j \in N\backslash\{i\}$, $g_{ij} = 1$ when i maintains a link with[6] j; $g_{ij} = 0$ when there is no link from i to j. Let $X = \{0, 1\}$ and $G_i = \{0, 1\}^{n-1}$ represent the respective space of volunteering and linking strategies of player i. Let $S_i = X \times G_i$ be the entire strategy space of player i and $S = S_1 \times S_2 \times \cdots \times S_n$ be that of all players. A strategy profile $s = (x, g) \in S$ specifies the actions $x = (x_1, x_2, \ldots, x_n)$ and the network of relations $g = (g_1, g_2, \ldots, g_n)$, where $g_i = (g_{i1}, ..., g_{ii-1}, g_{ii+1}, ..., g_{in}), \forall i \in N$. We focus on pure strategies throughout the paper.

By definition, the network of relations g is a directed graph. Let $N_i^{in}(g) = \{j \in N : g_{ji} = 1\}$ be the set of players forming a link to i and $N_i^{out}(g) = \{j \in N : g_{ij} = 1\}$ be the set of those receiving a link from i. Define the closure of g as an undirected graph, denoted by $\bar{g} = cl(g)$, where $\bar{g}_{ij} = \max\{g_{ij}, g_{ji}\}, \forall i, j \in N$. We call $j \in N_i(\bar{g}) = \{j \in N : \bar{g}_{ij} = 1\}$

[6]Throughout the text, we use “maintain a link with”, “form a link with” and “link with” interchangeably.

as a *neighbor* of player i. Players i and j are *linked* if $\bar{g}_{ij} = 1$ and are *minimally linked* if they are linked and $g_{ij}g_{ji} = 0$.

Moreover, we consider both *one-way* and *two-way flow of benefits* in networks. In the former case, a link that player i forms with j only allows i to access the public good provided by j; while in the latter case, the link allows both players to access the public good provided by the other. Meanwhile, a volunteer enjoys social benefits when her provision of the public good reaches others. Then the one-way (respectively, two-way) transmission of the public good leads to one-way (respectively, two-way) flow of social benefits.[7] We assume that the value of a volunteer's social benefits is increasing and concave in the number of recipients of her provision. For example, in a controlled lab experiment, [4] finds that altruism is a partially congested good: doubling the number of recipients increases but does not double the warm-glow utility to the giver.

Therefore, when the flow of benefits is one-way, the payoff to player i from strategy profile $s = (x, g)$ is:

$$\Pi_i^{ow}(s) = f(x_i + \sum_{j \in N_i^{out}(g)} x_j) - cx_i - k|N_i^{out}(g)| + b_i(|N_i^{in}(g)|)x_i, \tag{2.1}$$

where $f(\cdot)$ represents the benefits of consuming the public good, $c > 0$ is the cost of producing the public good, $k > 0$ is the cost of maintaining a link and $b_i(\cdot)$ denotes player i's social benefits from volunteering. Throughout the text, $|\mathcal{A}|$ represents the cardinality of a set $\mathcal{A}$. When the flow of benefits is two-way, the payoff is:

$$\Pi_i^{tw}(s) = f(x_i + \sum_{j \in N_i(g)} x_j) - cx_i - k|N_i^{out}(g)| + b_i(|N_i(g)|)x_i. \tag{2.2}$$

By abuse of notations, $\forall y \in \mathbb{N}$, write $f'(y) \equiv f(y+1) - f(y)$ and $f''(y) \equiv f'(y+1) - f'(y)$. To focus on interesting scenarios, we assume that $\forall y \in \mathbb{N}$, $f'(y) > 0$, $f''(y) < 0$, $f(0) = 0$, $f'(0) > c > k > 0$ and $c - k$ is sufficiently large. Then there exist $\check{y} >$

[7] See payoff functions below for the exact formulation.

$\hat{y} \geq 1$ such that $\hat{y} \equiv \arg\max_{y \in \mathbb{N}}[f(y) - cy] = \max\{y \in \mathbb{N} : f'(y-1) \geq c\}$ and $\check{y} \equiv \arg\max_{y \in \mathbb{N}}[f(y) - ky] = \max\{y \in \mathbb{N} : f'(y-1) \geq k\}$. Note that $\hat{y}$ (respectively, $\check{y}$) represents the optimal units of the public good that a player wants to access when the marginal access cost is c (respectively, k). Similarly, $\forall m \in \mathbb{N}$, write $b_i'(m) \equiv b_i(m+1) - b_i(m)$ and $b_i''(m) \equiv b_i'(m+1) - b_i'(m)$. We assume that $\forall m \in \mathbb{N}$, $b_i'(m) > 0$, $b_i''(m) < 0$, $b_i(0) = 0$, and $b_i'(m) > b_j'(m)$ if and only if $b_i'(0) > b_j'(0)$. Then there exist $\bar{m}_i > \tilde{m}_i \geq 1$ such that $\tilde{m}_i \equiv \arg\max_{m \in \mathbb{N}}[b_i(m) - km] = \max\{m \in \mathbb{N} : b_i'(m-1) \geq k\}$ and $\bar{m}_i \equiv \arg\max_{m \in \mathbb{N}}[f(1+m) + b_i(m) - km] = \max\{m \in \mathbb{N} : f'(m) + b_i'(m-1) \geq k\}$. Here, $\tilde{m}_i$ (respectively, $\bar{m}_i$) is the maximum number of outgoing links player i will form with free riders (respectively, volunteers) when she volunteers. In addition, we assume that $b_i(n-1) > c, \forall i \in N$. This allows us to focus on interesting scenarios in which the consideration of social incentives plays a role when player i decides whether or not to volunteer.

Let s_{-i} be the strategies of all players other than i. For either type of flow, define a strategy s_i as a *best response* of player i to s_{-i} if

$$\Pi_i(s_i, s_{-i}) \geq \Pi_i(s_i', s_{-i}), \forall s_i' \in S_i.$$

Let $BR_i(s_{-i})$ denote the set of player i's best responses to s_{-i}. Then a strategy profile $s^* = (s_1^*, \ldots, s_n^*)$ is a *Nash equilibrium (NE)* if $s_i^* \in BR_i(s_{-i}^*)$, $\forall i \in N$. A Nash equilibrium is *strict* if $\{s_i^*\} = BR_i(s_{-i}^*)$, $\forall i \in N$, *i.e.*, every player has a unique best response to others' strategies. Moreover, a strategy profile s *Pareto dominates* another profile $s' \neq s$ if $\Pi_i(s) \geq \Pi_i(s'), \forall i \in N$ and the inequality is strict for at least one player.

We now define some more concepts and notations to describe the equilibrium profiles. Let $V \equiv \{i \in N : x_i = 1\}$ be the set of volunteers and $F \equiv \{i \in N : x_j = 1\}$ be the set of free riders.

In addition, there is a path in $\bar{g}$ between players i and j if either $\bar{g}_{ij} = 1$ or there exist players $j_1, j_2, ..., j_m$ who are distinct from each other and from i and j, and satisfies $\bar{g}_{ij_1} = \bar{g}_{ij_2} = ... = \bar{g}_{ij_m} = 1$. We say a network $\bar{g}$ is connected if there exists a path between every pair of players; and $\bar{g}$ is minimally connected if it is connected and there is a unique path between every pair of players. We call an undirected network $\bar{g}$ as a simple graph if it is minimally connected, and a multigraph if it allows multiple links between two players. A *core-periphery network* consists of two groups of players, *hubs* N_{HUB} and *spokes* N_{SPK}. For any player $i \in N_{HUB}$, $N_i(\bar{g}) = N \backslash \{j\}$; for any player $j \in N_{SPK}$, $N_j(\bar{g}) = N_{HUB}$. $\bar{g}$ is a *core-periphery simple network* if it is both simple and has a core-periphery structure; $\bar{g}$ is a *core-periphery multi network* if it is both a multigraph and has a core-periphery structure. Moreover, we define an *independent set* of $\bar{g}$ as a subset of players such that no two players are linked. $\bar{g}$ is a *multipartite simple network* if it is a simple graph and the set of players can be partitioned into multiple independent sets. Note that a core-periphery simple network is a special case of a multipartite simple network, in which each hub player forms an independent set. Finally, we say that a network has a *nested split* structure if for any pair of players i and j, $N_j(\bar{g}) \subset N_i(\bar{g})$ whenever $|N_i(\bar{g})| > |N_j(\bar{g})|$.

2.4 Shape of Equilibrium Profiles

In this section, we characterize (strict) Nash equilibria for each type of network flow. We begin with some common elements for the two types. First, returns from consuming the public good, $f(\cdot)$, are increasing and concave. Second, the costs of personal provision, c, and of forming each additional link, k, are positive constants with $c > k$. Third, the public goods provided by oneself and the neighbors are substitutable. Recall that $\hat{y} = \arg\max_{n \in \mathbb{N}}[f(y) - cy]$ and $\check{y} = \arg\max_{n \in \mathbb{N}}[f(y) - ky]$, then we have the following.

Observation 2.1. *For either type of flow, the number of volunteers in every NE is at least $\hat{y}$.*

Observation 2.2. *For the mere sake of the consumption benefits of the public good, a player will only link with volunteers and will sponsor at most $\check{y}$ number of the links.*

Now we proceed to present the equilibrium profiles for the two types of flows.

2.4.1 One-way Flows

When the flow of benefits is one-way, only the sponsor of a link could access the public good provided by the other player involved. Therefore, volunteers would not be able to gain social benefits from actively linking with others. Then we have the following.

Proposition 2.1. *When the flow of benefits is one-way, every player in equilibrium only forms links with volunteers. In a strict NE, $\hat{y} \leq \sum_{i \in N} x_i^* \leq \check{y}$ and g^* is a core-periphery network in which every hub player is a volunteer receiving links from everyone else and every spoke player does not exert efforts.*

Figure 1 illustrates the strict NE for one-way flow of benefits.

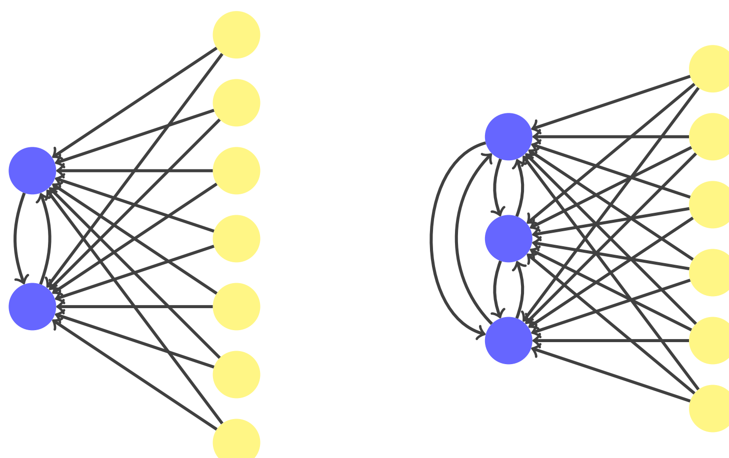

Figure 2.1 – Examples of strict Nash equilibria for one-way flows of benefits with $n = 9$, $\hat{y} = 2$ and $\check{y} = 3$. Blue (respectively, yellow) nodes represent volunteers (respectively, free riders).

Proof. We provide all proofs in Appendix A. □

The intuition of the results follows directly from the fact that for one-way flows, individuals form outgoing links merely for the sake of consumption benefits of the public good. If the total number of volunteers is at least $\check{y} + 1$, then every player will randomly pick $\check{y}$ (for a free rider) or $\check{y} - 1$ (for a volunteer) of them to link with, and nobody's best response is strict. This implies that in a strict NE, there are at most $\check{y}$ number of volunteers. Every volunteer's provision is pivotal to everyone else's consumption benefits and so receives incoming links from all other $n-1$ players. This results in a core-periphery multi network.

2.4.2 Two-way Flows

When the flow of benefits is two-way, a link transmits the public good provided by either player involved to the other, no matter who pays the linking cost. Therefore, while a free rider forms links just to access the public good, a volunteer's incentives for forming outgoing links come from both the consumption benefits of accessing others' provision of the public good and the social benefits of communicating her own provision to others. Moreover, a volunteer's social benefits are increasing and concave in the number of recipients of the public good she provides; and different recipients generate the same amount of social benefits, *ceteris paribus*. Therefore, there is still a fixed upper bound on the number of outgoing links a volunteer would like to form; she will prefer linking with another volunteer to linking with a free rider since the former also generates consumption benefits; and she will be indifferent among all volunteers (respectively, free riders) to link with.

Proposition 2.2. *When the flow of benefits is two-way, a volunteer in equilibrium may form links with free riders. A strict NE has a nested split network structure, which belongs to one of the following two types:*

- type I: $\hat{y} \leq \sum_{i \in N} x_i^* \leq \check{y}$, g^* *is a core-periphery simple network in which every hub player volunteers while individuals in the spoke do not exert efforts, and links between the hubs and the spokes could go in both directions;*
- type II: $\sum_{i \in N} x_i^* \geq \check{y} + 1$, g^* *is a multipartite simple network in which all links between the volunteers and the free riders go from the former to the latter; and there could be three groups of volunteers, say groups A, B and C: volunteers in groups A and B are minimally linked with each other and sponsor links with every group*

C volunteer, group A volunteers also form links with every free rider while group C volunteers do not sponsor any links.

Figure 2 illustrates the strict NE for two-way flow of benefits.

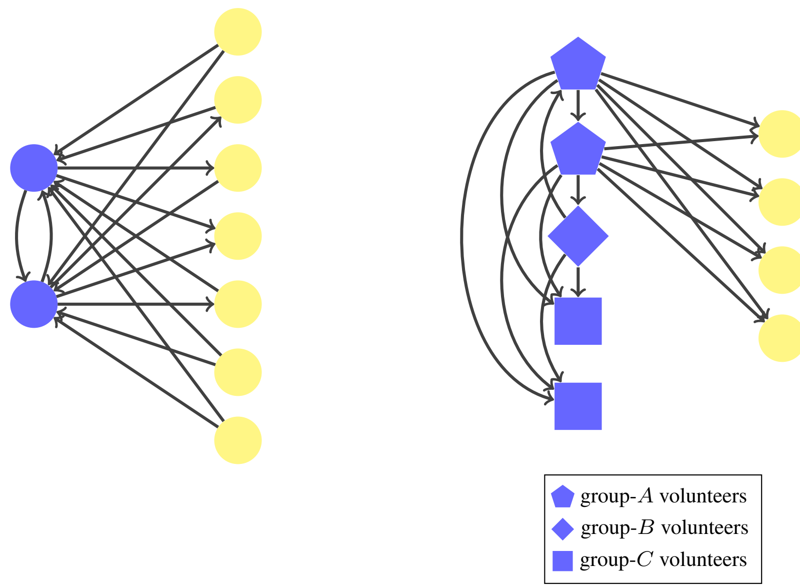

Figure 2.2 – Examples of strict Nash equilibria for two-way flows of benefits with $n = 9$, $\check{y} = 2$ and $(\tilde{m}_1, \tilde{m}_2, ..., \tilde{m}_9) = (50, 50, 5, 2, 2, 1, 1, 1, 1)$. Notes: i) every strict NE belongs to either type I or type II; ii) in every type-II strict NE, players with $\tilde{m}_i = 50$ are group-A volunteers and the one with $\tilde{m}_i = 5$ is a group-B volunteer; iii) type-I strict NEs with the same number of volunteers could differ in the directions of links between any pair of players while type-II strict NEs with the same number of volunteers could differ in the directions of links within group-A volunteers or between group-A and group-B volunteers; iii) for any volunteer i in a type-I strict NE, the number of sponsored outgoing links to free riders is no more than $\min\{\tilde{m}_i - |N_{i,V}|, |F|\}$, where $|N_{i,V}|$ is the number of volunteers as player i's neighbors and $|F|$ is the number of free riders; iv) for any volunteer i in a type-II strict NE, the number of sponsored outgoing links to free riders is either 0 (if $\tilde{m}_i - |N_{i,V}| \leq |F| - 1$) or $|F|$ (otherwise).

We discuss the points to note underlying the results. First, whenever the number of volunteers in equilibrium is no more than $\check{y}$, a player could always gain in consumption benefits by linking to an unlinked volunteer. This means every volunteer's provision is pivotal to others, leading to a core-periphery simple network in which the core players are volunteers and the periphery do not exert volunteering efforts. Second, recall that a volunteer prefers linking with another volunteer than linking with a free rider, then i) a volunteer forms links with all other unlinked volunteers before linking with free riders;

ii) when the number of volunteers exceeds $\check{y}$ in a strict NE, it could not be more than $\max_{i \in N} \tilde{m}_i$ since otherwise, no volunteer would sponsor links with free riders. Third, when the number of volunteers exceeds $\check{y}$, a volunteer's provision is no longer pivotal to every free rider while a free rider is indifferent among the volunteers to form links with. This means in a strict NE, links between volunteers and free riders all go from the former to the latter. Fourth, since a volunteer is indifferent among the other volunteers (respectively, the free riders) to form links with, then in a strict NE with more than $\check{y}$ number of volunteers, volunteers that maintain outgoing links are minimally linked with each other. The remaining volunteers form an independent set. This results in a multipartite simple network in which there could be three groups of volunteers as described in the proposition.

2.4.3 The Law of the Few

One major finding of $G\&G$ is that their equilibrium results exhibit the empirically observed law of the few: in large societies, a small proportion of the population provide public goods. We check if after the consideration of social incentives, the equilibrium still has that property. The answer is yes and follows from Propositions 2.1 and 2.2.

Corollary 2.1. *For either one-way or two-way flow of benefits, every strict NE exhibits the law of the few, $i.e.$, $\frac{\sum_{i \in N} x_i^*}{n} \to 0$ as $n \to \infty$.*

We note that in a strict NE, a volunteer's provision of the public good may or may not matter for another player's consumption benefits. The former happens for strict NE of one-way flows and type-I strict NE of two-way flows, and the total number of volunteers shall be no more than the maximum number of links a player will sponsor for consumption benefits only, which is $\check{y}$ by our notation. Then the conclusion follows directly. The latter happens for type-II strict NE of two-way flows. Since a free rider is indifferent among the volunteers to link with, then in a strict NE, it must be that volunteers form links with free riders. Moreover, since a volunteer i forms links with other volunteers first and will not link with free riders when the number of neighbors exceeds $\tilde{m}_i$, then whenever every free rider has access to the public good, the number of volunteers shall not exceed $\max_{i \in N} \tilde{m}_i$, and this number shall be no less than $n-1$. Since $\max_{i \in N} \tilde{m}_i$ is a constant, then the condition could not hold when $n \to \infty$. This implies that in large populations, every strict NE of two-way flows belongs to type I and the law of the few follows.

2.5 Effects of Social Incentives

It is tempting to think that the presence of social incentives will induce more volunteers in equilibrium. In this section, we first check if this is always the case. Relatedly, we check who will be the volunteers, or equivalently, if an individual choosing to free ride is due to enjoying lower social benefits from each recipient than a volunteer does. After these, we examine if the presence of social incentives will deliver higher efficiency.

To start with, we observe the following for the shapes of NE in the absence of social incentives.

Lemma 2.1. *For either one-way or two-way flow of benefits, in the absence of social incentives, in every strict NE, $\sum_{i \in N} x_i^* = \hat{y}$, and g^* has a core-periphery structure in which every player in the hub volunteers and receives links from everyone else and every player in the spoke does not exert volunteering efforts.*

Without social incentives, for either type of flow, a player makes volunteering decision merely based on the consideration of consumption benefits of the public good. Recall that the marginal consumption benefits exceed (fall below) the personal provision cost whenever the amount accessed is less (more) than $\hat{y}$ units, and it is cheaper to access a unit of the public good by linking with others than by exerting one's own efforts. Therefore, a player will choose to volunteer (free ride) when less (more) than $\hat{y}$ number of others do so. This means there are $\hat{y}$ number of volunteers in every equilibrium.

2.5.1 Do Social Incentives Always Induce More Volunteers and Who Are the Volunteers?

Lemma 2.1 implies that if the presence of social incentives brings more volunteers, then there are at least $\hat{y}+1$ number of them in equilibrium. We check if this holds in every strict NE. In addition, we check whether volunteers are always those with higher social benefits from each recipient.

Proposition 2.3. *When the flow of benefits is one-way, i) the presence of social incentives does not necessarily lead to more volunteers in every strict NE; ii) a volunteer not necessarily enjoys higher social benefits from each recipient than a free rider does in a strict NE.*

Proposition 2.4. *When the flow of benefits is two-way, i) the presence of social incentives will lead to more volunteers in every strict NE if and only if at least $\hat{y}+1$ number of players enjoy sufficiently high social benefits from each recipient; ii) a volunteer not necessarily enjoys higher social benefits from each recipient than a free rider does in a strict NE.*

The intuition is as follows. Based on the notion of strict NE, a strategy profile is stable if nobody has an incentive to deviate. For one-way flows, a player's social benefit from volunteering is a function of the number of incoming links she receives, which only depends on others' linking decisions. Then given others' strategies, whenever that number is zero, an increase in her own social benefits from each recipient will not affect her payoff from volunteering and thus will not change her volunteering decision. Since a free rider in any strict NE of one-way flows does not have incoming links, then for the strategy profile with $\hat{y}$ number of volunteers, an increase in marginal social incentives will not cause any free rider to deviate. Hence, for one-way flows, no matter how high the marginal social incentives are, it is always possible for a strict NE to have the same number of volunteers as when the incentives are absent. In addition, players with higher marginal social incentives may choose to free ride due to zero incoming link.

For two-way flows, a player's social incentives come from both incoming and outgoing links. Then given others' strategies, when her social benefits from each recipient increase, she may form more outgoing links, which will in turn raise her social benefits from volunteering. Therefore, when the marginal social incentives are sufficiently high for at least $\hat{y}+1$ number of the players, there will be more than $\hat{y}$ number of volunteers in every strict NE. On the other hand, since a player's social benefits still depend on the number of incoming links, then the payoff of volunteering will be lower when that number is smaller. Recall that in type-I strict NE of two-way flows, link directions are not fixed. And in type-II strict NE, whenever group B and C volunteers exist simultaneously, a group C volunteer always has less incoming links than a free rider does. Therefore, in these cases, a player may choose to free ride just due to too few incoming links, rather than low social benefits from each recipient.

The above analysis shows that in general, we could not conclude that in every strict NE, the presence of social incentives will lead to more volunteers and a volunteer enjoys higher social benefits from each recipient than a free rider does. This is because there exist many strict NE that differ in the numbers of volunteers, the identities of volunteers for a given number of them or the directions of links, which is related to an equilibrium selection problem.

In the rest of this subsection, we adopt a solution concept called *stochastically stable equilibrium*[8] for equilibrium selection and further analysis of the impacts of social incentives. Individuals are considered to be boundedly rational and make decisions myopically and periodically, due to say mistakes, mutations or limited information about the game. More formally, at each discrete point of time, one player is randomly selected to update her strategy based on the current play of others. The update follows a logit quantal response rule: a player chooses best reply with high probability and chooses other strategies with some non-zero small probabilities. We assume that a mutation is less likely if it entails a larger loss in payoff. A strategy profile is a *stochastically stable equilibrium (SSE)* if it

[8]For an introduction of this solution concept, see among others, [29], [36] and [54].

has a positive probability in the limit distribution when the sizes of stochastic perturbations become vanishingly small. More details of the dynamic setup are available in Appendix B. The introduction of perturbed best-response dynamics allows us to study the relative robustness of different strict NE. Below we make an attempt to characterize the conditions under which SSE involve more volunteers in the presence of social incentives, who enjoy higher social benefits from each recipient than free riders.

Remark 2.1. *In SSE, the presence of social incentives will lead to more volunteers and the volunteers are those who enjoy higher social benefits from each recipient of their provision if there exist a sufficient number of individuals with high enough social benefits, which are also significantly larger than the rest of the population.*

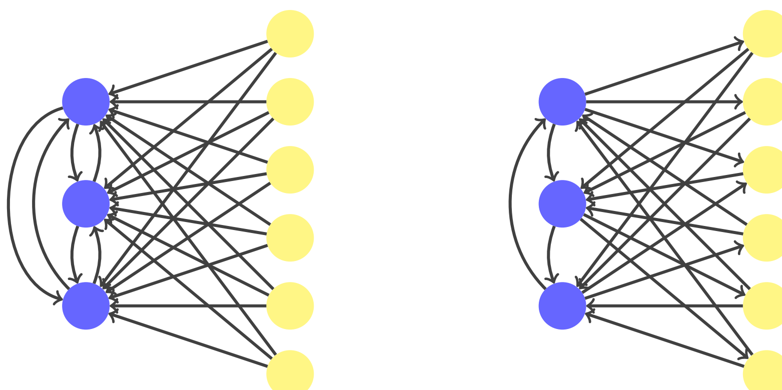

Figure 2.3 – Examples of stochastically stable equilibria for one-way flow of benefits (the left graph) and two-way flow of benefits (the right graph) with $n = 9$, $\hat{y} = 2$, $\check{y} = 3$ and $(\tilde{m}_1, \tilde{m}_2, ..., \tilde{m}_9) = (50, 50, 50, 1, 1, 1, 1, 1, 1)$. Notes: i) the left graph represents the unique SSE for one-way flow of benefits in which players with $\tilde{m}_i = 50$ are the volunteers; ii) every SSE for two-way flow of benefits is a type-I strict NE in which the three players with $\tilde{m}_i = 50$ are the volunteers, and different SSE vary in terms of link directions between the players.

The rationale for this remark is as follows. First, as mentioned above, the shape of SSE shows the relative stability of different strategy profiles, which depends on the probability and thus the cost of a transition between every pair of states. We note that if due to some exogenous changes, the cost of every transition into a state is lower than that of every transition out of it, then that state is the unique SSE. Second, as is known in the literature, only the cheapest route matters for calculating the cost of a transition. For either type of flow, there are typically three candidate cheapest routes for a transition that involves the birth (death) of a volunteer. Route 1 involves that player's own mistake of starting (stopping) volunteering; route 2 involves the player's best response to others' mistakes in adding (severing) links to her; route 3 involves the player's best response to others' mistakes in stopping (starting) volunteering. We note that since a player only enjoys social benefits when she volunteers, then the cost of the birth (death) of that player using route 1 or 2 is declining (increasing) in her social benefits from each recipient, while that using route 3 is related to others' social benefits from each recipient. When her social benefits from each recipient are high enough, which are also sufficiently larger than others', the cheapest route belongs to one of the first two routes and the cost of every transition involving the birth of that volunteer is lower than that of transition involving the death of the volunteer. Therefore, when there exist a sufficient number of players with high enough social benefits from each recipient, which are also much larger than the rest of the population, SSE involves those players with high social benefits from each recipient as the volunteers.

2.5.2 Do Social Incentives Make Everyone Better Off?

In this subsection, we set out to examine the welfare effects of social incentives. We adopt the concept of Pareto dominance when comparing the efficiency of two strategy profiles. We first examine whether the strict NE are Pareto ranked for each type of flow.

Proposition 2.5. *When the flow of benefits is one-way, a strict NE with more volunteers Pareto dominates another with less volunteers.*

Proposition 2.6. *When the flow of benefits is two-way, i) a type-I strict NE with more volunteers Pareto dominates another with less volunteers; ii) a type-II strict NE with more volunteers and a type-I strict NE with less volunteers do not Pareto dominate each other in general.*

The above results show that for either type of flow, whenever the equilibrium network has a core-periphery structure and the number of volunteers is not large, having more volunteers in equilibrium is always beneficial to everyone. When the network flow is of two-way, having many volunteers in a type-II equilibrium may not make everyone better off than in a type-I equilibrium with less volunteers. This is because in the case of overabundance, volunteers would have to form links to every free rider. They may then receive less payoffs than in an equilibrium in which they get access to similar amount of public goods and social benefits, but have much less costly links to form.

These results together with the ones in Section 5.1 lead to the following.

Corollary 2.2. *When the flow of benefits is one-way, social incentives not necessarily make every Nash player better off, due to possible coordination failure.*

Corollary 2.3. *When the flow of benefits is two-way, social incentives not necessarily make every Nash player better off, due to possible overabundance of volunteers and costly linking.*

We further note that for two-way flows, if the cost of linking is significantly low relative to the social benefits from each recipient of the public goods, then everyone will be better off with the consideration of social incentives. Then we have the following.

Corollary 2.4. *For either type of network flow, stochastic stability based on logit quantal response dynamics predicts that every player will be better off with the consideration of social incentives if the social benefits from each recipient of the public goods are significantly large relative to the linking cost for sufficiently many individuals.*

2.6 Extensions

2.6.1 Heterogeneity in Consumption Benefits and Production Costs of the Public Good as well as Linking Costs

In the basic model, players are only heterogeneous in their valuations of the social incentives. In reality, heterogeneity could also lie in consumption benefits and production costs of the public good as well as linking costs. (For examples, see among others, [9], [5], [34] and [37].) If we take the additional heterogeneity into account, then the payoff functions for player i with strategy profile $s = (x, g)$ become:

$$\Pi_i^{ow}(s) = f_i(x_i + \sum_{j \in N_i^{out}(g)} x_j) - c_i x_i - k_i |N_i^{out}(g)| + b_i(|N_i^{in}(g)|) x_i, \tag{2.3}$$

$$\Pi_i^{tw}(s) = f_i(x_i + \sum_{j \in N_i(g)} x_j) - c_i x_i - k_i |N_i^{out}(g)| + b_i(|N_i(g)|) x_i. \tag{2.4}$$

Similar to previous definitions, let $\hat{y}_i \equiv \arg\max_y [f_i(y) - c_i y]$, $\tilde{y}_i \equiv \arg\max_y [f_i(y) - k_i y]$. To focus on interesting cases, we still assume that $c_i > k_i > 0, \forall i$, and so $\hat{y}_i \leq \tilde{y}_i, \forall i$.

Proposition 2.7. *For one-way flow of benefits with generalized heterogeneity, in every strict NE* $s^* = (x^*, g^*)$, $\min_{i \in N} \hat{y}_i \leq \sum_{i \in N} x_i^* \leq \min_{i \in N} \tilde{y}_i$ *and* g^* *is a core-periphery network in which every hub player volunteers and receives links from everyone else while every spoke player does not exert efforts.*

Proposition 2.8. *For two-way flow of benefits with generalized heterogeneity, a strict NE* $s^* = (x^*, g^*)$ *has the following shapes. i) When* $\sum_{i \in N} x_i^* \leq \min_{i \in N} \tilde{y}_i$*,* g^* *has a core-periphery structure in which volunteers lie at the core and links between free riders and volunteers go in either direction. ii) When* $\sum_{i \in N} x_i^* \geq \min_{i \in F} \tilde{y}_i$*,* g^* *is a multipartite graph in which links between volunteers and free riders always go from the former to the latter, and free riders (and some of the volunteers) form an independent set. iii) When* $\min_{i \in V} \tilde{y}_i \leq \sum_{i \in N} x_i^* \leq \min_{i \in F} \tilde{y}_i$*,* g^* *is a multipartite graph in which every volunteer is minimally linked with all free riders while volunteers may or may not be minimally linked with each other.*

We note that the first two are the counterparts of the shapes identified as type I and type II in Proposition 2.2. The third one exists only when the condition $\min_{i \in V} \tilde{y}_i \leq \sum_{i \in N} x_i^* \leq \min_{i \in F} \tilde{y}_i$ is satisfied. Therefore, every strict NE in the basic model is still a strict NE here.

2.6.2 Imperfect Substitutability Between One's Own and Others' Volunteering Efforts

As mentioned in [14], an individual's own volunteering effort could be more beneficial to herself than to others. This may be because some public good is lost when transmitted to neighbors, say due to imperfect communication or tacit knowledge [37]. We now take such situations into consideration. Let $\delta \in (0,1]$ measure the extent to which own efforts and neighbor's efforts are substitutes. Then the payoff functions become

$$\Pi_i^{ow}(s) = f(x_i + \delta \sum_{j \in N_i^{out}(g)} x_j) - cx_i - k|N_i^{out}(g)| + b_i(|N_i^{in}(g)|)x_i, \tag{2.5}$$

$$\Pi_i^{tw}(s) = f(x_i + \delta \sum_{j \in N_i(g)} x_j) - cx_i - k|N_i^{out}(g)| + b_i(|N_i(g)|)x_i. \tag{2.6}$$

In this case, the effective cost of obtaining one unit of public good from neighbors is $k_\delta \equiv k/\delta$. To focus on interesting scenarios, we still assume that $c > k_\delta$. In addition, let $\tilde{y}_\delta \equiv \arg\max_y [f(y) - k_\delta y]$.

Proposition 2.9. *For one-way flow of benefits with the public good's production function as $f(x_i + \delta \sum_{j \in N_i^{out}(g)} x_j)$ where $\delta \in (0,1]$, in every strict NE $s^* = (g^*, x^*)$, $\hat{y}/\delta \leq \sum_{i \in N} x_i^* \leq \tilde{y}_\delta/\delta$ and g^* is a core-periphery network in which every hub player is a volunteer receiving links from everyone else and every spoke player does not exert efforts.*

Proposition 2.10. *For two-way flow of benefits with the public good's production function as $f(x_i + \delta \sum_{j \in N_i(g)} x_j)$ where $\delta \in (0,1]$, a strict NE $s^* = (x^*, g^*)$ has a nested split network structure, which belongs to one of the following two types:*

- type I': $\hat{y} \leq \sum_{i \in N} x_i^* \leq \tilde{y}_\delta/\delta$, *$g^*$ is a core-periphery network such that every hub player volunteers while individuals in the spoke do not exert efforts, and links between the hub and the spoke could go in both directions;*

- type II': $\sum_{i \in N} x_i^* \geq \tilde{y}_\delta/\delta + 1$, *$g^*$ is a multipartite network in which links go from* *volunteers to free riders and between volunteers, and free riders (and some of the volunteers) form independent sets.*

Therefore, our general characterization of the strict NE in the basic model is robust to this extension.

2.7 Conclusion

This paper examines the role of social incentives in public goods games in endogenous networks. We find that sufficiently large social benefits from each recipient of the public goods would induce more volunteers and active sharing of the public goods only when the network flow of benefits is two-way. Moreover, for either type of network flow, stochastic stability predicts that every player will be better off with the consideration of social incentives if such benefits per recipient are sufficiently large relative to the linking cost for sufficiently many individuals.

There are two potential extensions to make in future work. First, the decision on public good provision in this model is binary. One could further investigate the case in which individuals decide how much to produce and examine the effects of social incentives on both the extensive and intensive margins of public goods provision. The current setup focuses on the extensive margin. Second, our analysis of the effects of social incentives has mainly dealt with the quantity side, $i.e.$, the size of the set of volunteers. Future work could study the effects of different types of social incentives on both the quantity and quality of the public goods produced in endogenous networks.

CHAPTER 3

MULTIPLE CLIMATE CLUBS

3.1 Introduction

Strong free-riding incentives and the norm of voluntary participation have made it difficult to forge stable international agreements on climate change. [45] proposes the idea of a global climate club and shows that small trade penalties on non-participants can induce a large stable coalition. In this paper, we build upon his idea and consider the formation of multiple (overlapping) climate clubs. We study effects of the formation of multiple clubs, as well as the role of free trade agreements (FTAs) among club members therein.

More specifically, we consider a climate club game in an open economy with three stages. In the first stage, every country announces with which other countries it would like to form clubs. In the second stage, every country chooses tax rates on domestic carbon emissions from energy production and tariff rates on energy imports. In the third stage, production, trade and consumption take place. As mentioned above, we allow a country to join one or multiple clubs and consider regimes both with and without FTAs as parts of the club treaties.

The main findings are as follows. First, when club treaties do not involve FTAs, a fully cooperative structure in which all countries are (in)directedly connected to each other via overlapping clubs is stable against all possible deviations and efficient. This comes as a result of several facts when FTAs are absent: countries are symmetric in all structures with interior solutions; structures that are fully cooperative are equivalent in terms of policy making; and countries in a more connected/cooperative structure enjoy higher welfare as the benefits from lower trade barriers dominate the costs associated with higher carbon taxes. Therefore, every subgroup of countries has the same incentive to deviate to the most cooperative structures, which generate the highest global welfare.

Second, when club treaties involve FTAs, a structure with the most clubs could be stable against all possible deviations and efficient if the structures with interior solutions are those in which all countries are geometrically symmetric, or those in which countries are geometrically asymmetric and the density of clubs is not very high. Otherwise, a structure with less clubs may be stable but not efficient. To understand these results, we first note that in any parameter range, the structure with the most clubs is efficient. Next, for structures in which all countries are geometrically symmetric, stability is in line with efficiency because structures that are fully connected/cooperative are equivalent in terms of policy making and countries in a more connected/cooperative structure enjoy higher welfare. For structures in which countries are geometrically asymmetric, those with the same connectedness are no longer equivalent in terms of policy making and the density of the clubs matters. More specifically, in asymmetric structures with a not very high density of clubs, a country joining more clubs would be better off than others joining less in the same structure, as well as those in a less asymmetric structure. The reverse holds in asymmetric structures with a high density of clubs. Therefore, when a parameter range involves an asymmetric structure having interior solutions, which structure would be stable depends on which other structures also have interior solutions in the same range.

Third, the formation of multiple climate clubs improves stability of the equilibrium structures for emissions coefficients that are not large. More specifically, we find that when club treaties do not involve FTAs, the formation of multiple clubs generates more stable and efficient structures for emissions coefficients that are intermediate. When club treaties involve FTAs, the formation of multiple clubs allows structures that are stable and efficient to exist for small emissions coefficients; it also generates more stable though less efficient structures than the fully connected structure for emissions coefficients that are intermediate. The result in the absence of FTAs is due to the fact that when the emissions coefficient is sufficiently large, only partially-cooperative and non-cooperative structures have interior solutions. And the partially-cooperative structure, which only exists when multiple clubs are allowed, is more stable and efficient than the latter. The results in the presence of FTAs are due to the facts that when the emissions coefficient is small, the formation of multiple clubs allows structures with geometrically asymmetric countries and a not very high density of clubs to have interior solutions. When the emissions coefficient is intermediate and structures with geometrically asymmetric countries and a high density of clubs have interior solutions, a fully-connected structure which is the stable equilibrium if only one club is allowed to form, would no longer be stable as some of the countries would like to withdraw from clubs together and enjoy the higher welfare in the asymmetric structure.

Fourth, the inclusion of FTAs in club treaties enlarges the stable and efficient region of the equilibrium. More specifically, free trade among club members allows the grand coalition structure to be stable against all possible deviations and efficient whenever emissions coefficients are sufficiently large, which does not even have interior solutions in the absence of FTAs. When multiple clubs are formed, this region is further enlarged, and the stability of the equilibrium though not the efficiency may also be improved. In particular, structures with geometrically asymmetric countries, which are less connected/cooperative than the grand coalition, emerges in equilibrium. These structures are stable against all deviations and efficient for lower values of emissions coefficients than in the one-club case. For some intermediate values of emissions coefficients, these structures are more stable though less efficient than the grand coalition.

Finally, coalition-proofness and group stability may yield different equilibrium structures when multiple clubs are formed and FTAs are involved. More specifically, the formation of multiple clubs and the inclusion of FTAs would lead to differences in national welfare for countries that are geometrically asymmetric in a given structure. A group deviation may then fail to be self-enforcing if some of the countries find it profitable to deviate further.

3.1.1 Related Literature

This paper is related to the literature on international environmental agreements (IEAs). Many studies find that due to strong free-riding incentives, it is difficult to form stable coalitions on climate change with a large number of participants and/or large gains from cooperation (see, among others, [17], [7], [21] and [47]). This motivates a strand of literature to explore ways of reducing free riding incentives by linking climate change with international trade (see, among others, [24], [25], [26], [45], [22] and [38]). In particular, [45] proposes the idea of a global climate club and shows that small trade sanctions on non-participants could induce a large stable coalition. Relatedly, [22] and [38] find that the interaction between two instruments, emission taxes/caps and import tariffs, is essential for enhancing cooperation: the gains from free trade can provide strong incentives for countries to join a climate coalition. However, all of these models consider the formation of a single climate club. Recently, [48] and [52] demonstrate that structures with multiple overlapping climate clubs/coalitions could also be stable with a large number of participants, although no international trade is involved. This paper contributes to the literature by studying a model with multiple climate clubs in an open economy. We show that the formation of multiple clubs improves stability of the equilibrium, and free trade among club members enlarges the parameter range with stable and efficient equilibrium. In particular, equilibria that involve all countries as participants of the clubs, which may be more than three, could be stable against all possible deviations. When club treaties involve FTAs and the emissions coefficients are sufficiently large, the grand coalition would be the only stable and efficient structure.

In addition, as the stability of IEAs requires not only individual rationality but also coalitional and collective rationality (see, among others, [45]), this paper considers stability against profitable bilateral, multilateral and self-enforcing deviations. More specifically, we adopt bilateral and group stability from [27]), which rules out all profitable deviations by subgroups in a network. Since some of these deviations may not be self-enforcing, we also adopt the concept of coalition-proof Nash equilibrium proposed by [12]. To the best of our knowledge, except [45], [49] and [48], we are among the first ones to apply these coalitional stability concepts to the study of IEAs. Moreover, we compare the results generated from the various stability concepts, and find that equilibrium structures based on coalition-proofness and group stability could be different when multiple clubs are formed and FTAs are included.

The organization of the chapter is as follows. Section 3.2 presents the basic model setup for an economy with three nations. Section 3.3 analyzes the equilibrium of the basic model. Section 3.4 extends the model to an economy with four nations and studies its equilibrium. Section 3.5 summarizes the findings.

3.2 Basic Model

Consider a climate club game in an open economy with a set $N = \{1, 2, 3\}$ of symmetric countries. There are two goods: a numeraire good and an energy good.

In each country $i \in N$, a representative consumer chooses the amounts of the numeraire and energy goods to purchase, denoted by $\{x_i^D, e_i^D\}$, and solves

$$\max_{\{x_i^D, e_i^D\}} U_i(x_i, e_i^D) = x_i^D + \alpha e_i^D - \frac{\beta}{2}(e_i^D)^2$$
$$\text{subject to } x_i^D + p_i e_i^D = y_i \tag{3.1}$$

where p_i is price of the energy good, y_i is initial endowment of the numeraire good, and $\alpha, \beta > 0$ are parameters. We assume that consumers are price takers, and the price of the numeraire good is equal to one everywhere.

On the supply side in each country, both goods are produced by a representative firm with constant-returns-to-scale technologies. For the numeraire good, the market is perfectly competitive and the production does not emit pollution. Each country's market for the energy good is oligopolistic, served by both domestic and foreign firms. Let e_{ij}^S be the amount of the energy good supplied by firm i in market j. Then the market-clearing condition is

$$e_i^D = \sum_j e_{ij}^S \tag{3.2}$$

Meanwhile, we assume that the production of one unit of energy good generates θ units of pollution in the home country. A firm faces domestic carbon taxes on the emissions, as well as trade tariffs when exporting to foreign countries. Therefore, firm i solves

$$\max_{\{e_{ij}^S\}_{j=1,2,3}} \Pi_i(e_{ij}^S) = \sum_j [(\frac{p_j}{1+\tau_{ij}} - c - t_i\theta)e_{ij}^S] \tag{3.3}$$

where c is the constant marginal cost, t_i is the carbon tax on per unit of emissions, and τ_{ij} is the ad valorem tariff rate imposed by country j on products from i.

The government of each country determines the carbon tax and trade tariff rates mentioned above, as well as with which other countries to form climate clubs. Countries not in clubs will set both policies to maximize national welfare. For those in the same club, they will set the environmental policies to maximize joint welfare. Regarding the trade policies, we consider two regimes: with and without free trade agreements (FTAs) among club members, so as to examine the effects of linking trade policies with environmental ones in an economy that involves both international trade and (multiple) endogenous climate clubs. In the presence of FTAs, member countries will trade freely with each other, and set trade tariffs on non-members to maximize club welfare. Here, country i's welfare function is given by

$$\begin{aligned} W_i &= CS_i + \Pi_i + TR_i + CT_i - H_i \\ &= \frac{\beta}{2}(e_i^D)^2 + (p_i - c)e_{ii}^S + \sum_{j\neq i}[(\frac{p_j}{1+\tau_{ij}} - c)e_{ij}^S - \frac{p_i\tau_{ji}}{1+\tau_{ji}}e_{ji}^S] - \frac{h}{2}(\theta\sum_i\sum_j e_{ij}^S)^2 \end{aligned} \tag{3.4}$$

where $CS_i = \frac{\beta}{2}(e_i^D)^2$ is the consumer surplus, $TR_i = \sum_j[(p_i - \frac{p_i}{1+\tau_{ji}})e_{ji}^S]$ is the tariff revenue, $CT_i = t_i\theta \sum_j e^S_{ij}$ is the carbon tax and $H_i = \frac{h}{2}(\theta \sum_i \sum_j e^S_{ij})^2$ is the damage caused by the emissions (with h as the social cost of carbon). We further assume that an importing country will impose the same tariff rate on all sourcing countries that do not have FTAs with it. That is, $\tau_{ji} \equiv \tau_i$, for all j such that $\tau_{ji} > 0$.

The timing of the game is as follows. In the first stage, every country simultaneously announces with which other countries it would like to form climate clubs. Let σ_i denote country i's announcement and Ω_i be its strategy space. Then $\Omega_i = \{\{\emptyset, \emptyset\}, \{j, \emptyset\}, \{\emptyset, k\}, \{j, k\}\}$, where $\{\emptyset, \emptyset\}$ means no climate club to form; $\{j, \emptyset\}$ means a bilateral club with country j only; $\{\emptyset, k\}$ means a bilateral club with country k only; and $\{j, k\}$ means bilateral clubs with countries j and k respectively. In the second stage, given the clubs formed, countries determine optimal carbon taxes and trade tariffs. In the third stage, production, international trade and consumption take place.

3.3 Equilibrium Analysis

We now proceed to study the equilibrium of the basic model by analyzing backwards.

3.3.1 Stage Three

In each country $i \in N$, solving the optimization problem on the demand side gives the inverse demand function

$$p_i = \alpha - \beta e^D_i \tag{3.5}$$

Solving firm i's optimization problem gives the amount of energy goods it will supply in each market j

$$e^S_{ij} = \frac{\alpha + \sum_k [(c + t_k\theta)(1 + \tau_{kj})] - (n+1)(c + t_i\theta)(1 + \tau_{ij})}{(n+1)\beta} \equiv e^S_{ij}(t, \tau) \tag{3.6}$$

where $t = (t_1, t_2, t_3)$ and $\tau = (\tau_1, \tau_2, \tau_3)$ are vectors of carbon tax and trade tariff rates, respectively.

Substituting equation (6) into (2), (4) and (5), we obtain the equilibrium expressions for demand, prices and welfare

$$e_i^D \equiv e_i^D(t, \tau) \tag{3.7}$$

$$p_i \equiv p_i(t, \tau) \tag{3.8}$$

$$W_i \equiv W_i(t, \tau) \tag{3.9}$$

3.3.2 Stage Two

In this stage, given the climate clubs formed, each country government sets the environmental and trade policies. Figure 1 illustrates the possible club structures yielded from the first stage.

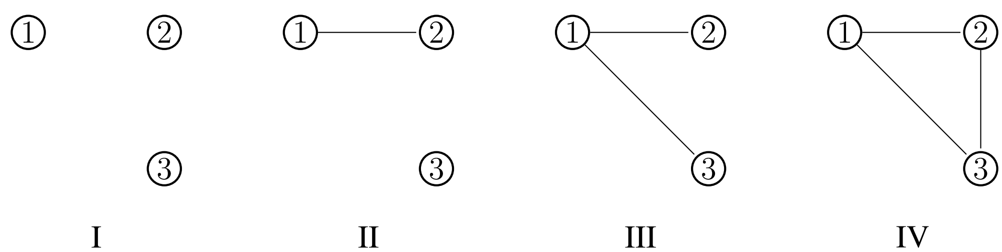

Figure 3.1 – Possible club structures resulted from the first stage ($n = 3$).

Since all countries have the same welfare function yielded from the third stage, it is then easy to see the following.

Observation 3.1. *In club structures I and IV respectively, countries are symmetric in policy decision making since these club structures preserve the symmetry.*

Moreover, note that no matter the climate clubs involve FTAs or not, countries in the same club always cooperate on environmental policies. That is, for each country i in club C_l,

$$t_i = t_{C_l} = \arg\max_t \sum_{j \in C_l} W_j(t, \tau)$$
$$\text{subject to } t_{C_l} \leq t_{C_{l'}}, \forall C_l, C_{l'} \text{ s.t. } C_l \cap C_{l'} \supseteq \{i\} \tag{3.10}$$

Here, the constraint applies to the situation in which country i joins club(s) in addition to C_l. The next observation follows immediately.

Observation 3.2. *When there are overlapping climate clubs, the carbon tax rates in all member countries of these clubs will be the same.*

Regarding trade policies, in the absence of FTAs, for each country i in club C_l,

$$\tau_i = \arg\max_{\tau} W_i(t, \tau) \tag{3.11}$$

In the presence of FTAs, for each country i in club C_l and for all $j \neq i$,

$$\tau_{ji} = \begin{cases} \arg\max_{\tau} \sum_{k \in C_l} W_k(t, \tau), & \text{if } i \text{ and } j \text{ do not belong to a common club} \\ 0, & \text{otherwise} \end{cases} \tag{3.12}$$

Based on these decision rules, for the case without FTAs, we have one more useful observation. In that case, for club structures III and IV, carbon taxes are set to maximize bilateral welfare everywhere, and trade tariffs are set to maximize national welfare everywhere. This symmetry in policies plus symmetric welfare functions from the last stage lead to complete symmetry between the two structures.

Observation 3.3. *In the absence of FTAs, club structures III and IV are identical.*

Simulation Results

To proceed further, we resort to numerical simulations to calculate the equilibria. We will present welfare results for $\alpha = \beta = 1, h = 0.01$, varying the remaining two parameters $c \in [0.1, 0.9]$ with step size 0.2, and $\theta \in [0.01, 10]$ with step size 0.01.[1]

Table B.1 (respectively, Table B.2) provides rankings of national welfare in the absence (respectively, presence) of FTAs. We refer the readers to Appendix for all tables.

[1] We consider smaller step size for θ as preliminary trials indicate that the equilibrium results are more sensitive to θ than to c.

Result 3.1. *In the absence of FTAs, the rankings of equilibrium national welfare across club structures for $n = 3$ are as listed in Table B.1. In particular, (part of) the following ranking holds:* $w_i^I < w_i^{III} = w_i^{IV}, \forall i$.

[Insert Table B.1 here.]

The results show that structures I, III and IV have interior solutions in the absence of FTAs. Regarding the welfare ranking, we see that the equalities are in accordance with Observations 3.1 and 3.3. For structures I and III, the internalization of environmental externalities by clubs raises carbon taxes, and subsequently reduces trade tariffs among the welfare-maximizing nations. The inequality then implies that cooperative environmental policies would make every country better off, which is the overall effect of higher carbon taxes and lower trade tariffs.

Result 3.2. *In the presence of FTAs, the rankings of equilibrium national welfare across club structures for $n = 3$ are as listed in Table B.2. In particular, (part of) the following two rankings hold:* $w_2^{III} < w_1^{IV} < w_1^{III}$ *and* $w_3^{II} < w_1^I < w_1^{IV} < w_1^{II}$.

[Insert Table B.2 here.]

For the first ranking, we know that every country in structure III (IV) sets the same carbon tax to maximize bilateral welfare. Besides, countries in structure IV are symmetric and the hub country in III has FTAs with the spokes. While carbon taxes are all set to maximize bilateral welfare, the unique advantage of enjoying free trade with both spokes makes the hub in III better off than not only a spoke in III but also a symmetric country in IV. For the second ranking, we know that countries in structure I (IV) are symmetric. Although forming a club makes every country in IV bear higher carbon taxes, the overall effect is dominated by the benefits of being able to free trade with each other. On the other hand, countries in structure II are asymmetric. The formation of a bilateral club leads to higher carbon tax rates in the member countries, but at the same time allows them to free trade with each other and charge higher tariffs on the non-member country. The overall effect of the asymmetric structure in II is again dominated by the trade effect, and the welfare of a member country will be higher than that of not only a symmetrically non-cooperative country in I but also a symmetrically cooperative country in IV.

From the above analysis, we see that forming a club has two effects on a country's policies: 1) a member country will set higher carbon tax (environmental effect); and 2) a member country will trade freely with others in the same club (trade effect in the presence of FTAs). The overall effect determines the rankings of equilibrium welfare.

3.3.3 Stage One

In this stage, every country announces the names of the other countries it would like to form climate clubs with. With the welfare rankings at hand, we derive the equilibrium club structures based on different notions of stability.

Definition 3.1. *A* profitable deviation *by $N' \subseteq N$ is a deviation that allows players in N' to simultaneously add absent links within N' and delete any link incident to at least one vertex in N', and strictly improves the expected payoff of each member of N'.*[2]

Definition 3.2. *An announcement strategy profile $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ is a* Nash equilibrium (NE) *if there is no profitable deviation for any player $i \in N$.*

Definition 3.3. *An announcement strategy profile $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ is a* bilaterally stable (BS) *if there is no profitable deviation for any $N' \subseteq N$ with $|N'| \leq 2$.*[3]

Definition 3.4. *An announcement strategy profile $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ is a* group stable (GS) *if there is no profitable deviation for any $N' \subseteq N$.*[4]

Definition 3.5. *An agreement is* self-enforcing *if and only if there is no $N' \subseteq N$ such that taking the actions of its complement as fixed, can agree to deviate in a way that makes all*

[2] See, for example, [27].
[3] See, for example, citeerol2018network.
[4] See, for example, [27].

of its members better off.[5]

Definition 3.6. *An announcement strategy profile* $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ *is a* coalition-proof Nash equilibrium (CPNE) *if there is no self-enforcing deviation for any* $N' \subseteq N$.[6]

Definition 3.7. *An announcement strategy profile* $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ *is* efficient *if it maximizes global welfare.*

Based on the above definitions, it is easy to see that a structure is NE as long as it is not profitable for any country to withdraw from club(s) unilaterally. When two (or more) countries want to form clubs and/or withdraw from clubs together, the structure is not BS (GS). Moreover, any profitable deviation by one or two players is self-enforcing. Therefore, any CPNE must be BS.

Table B.3 (respectively, Table B.4) summarizes the equilibrium structures according to the above stability and efficiency concepts for the trade regime without FTAs (respectively, with FTAs).

Result 3.3. *In the absence of FTAs, the equilibrium structures for* $n = 3$ *are as listed in Table B.3.*

Result 3.4. *In the presence of FTAs, the equilibrium structures for* $n = 3$ *are as listed in Table B.4.*

Proof. We provide all proofs in Appendix. □

From Result 3.1, we know that in the absence of FTAs, countries in structures I, III (IV) are symmetric, and structure III (IV) generates higher welfare for each country than I does. Therefore, when structures I, III and IV all have interior solutions, I is not GS or CPNE since all countries would deviate together by forming links among each other.

[5] See, for example, [12].
[6] See, for example, [12].

From Result 3.2, in the presence of FTAs, when two structures that preserve the symmetry (i.e., structures I and III) have interior solutions, the one with more clubs (i.e., structure III) generates higher national welfare for each country, which implies that the non-cooperative one is neither GS nor CPNE. When both symmetric and asymmetric structures have interior solutions (i.e., II and IV; III and IV; I, II and IV), the asymmetry allows the country(ies) joining more clubs to enjoy higher welfare than being in a symmetric structure with either more or less clubs, and the opposite holds for those joining less clubs. Since $n = 3$, this implies that an asymmetric structure in which two of the countries enjoy higher welfare (i.e., structure II) will be immune to bilateral and group deviations, and thus will be BS, GS and CPNE. The asymmetric structure with one hub and two spokes (i.e., structure III) will not be stable as the two lower-welfare spokes will deviate together by forming a club between themselves.

3.4 Extension

The results so far are for an economy with three countries. In this section, we extend the model to a case with four symmetric countries engaging in the same economic activities as in section 3.3. As the set of possible structures resulted from Stage 3 becomes larger (shown in Figure 2), one could expect the results to be more general.

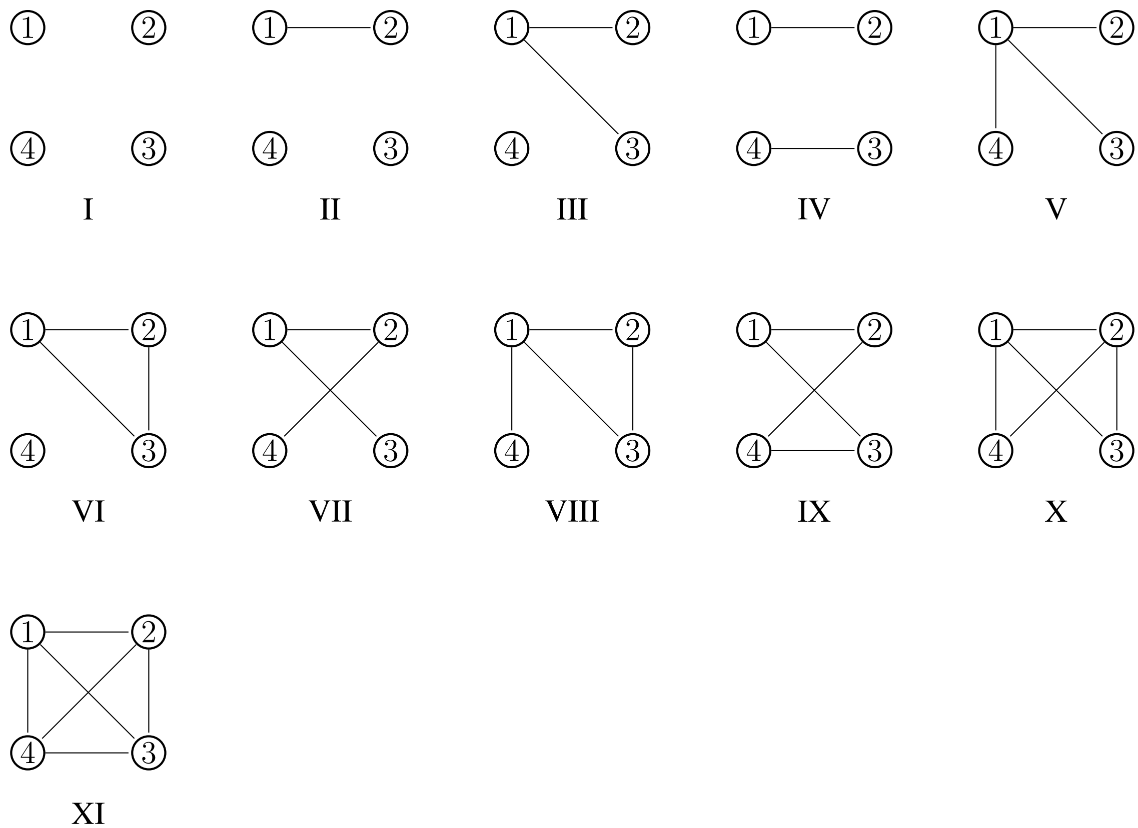

Figure 3.2 – Possible club structures resulted from the first stage ($n = 4$).

Result 3.5. *In the absence of FTAs, the rankings of equilibrium national welfare across club structures for $n = 4$ are as listed in Table B.5. In particular, structures in which all countries are (in)directly connected with each other generate the same welfare for all countries.*

Result 3.6. *In the presence of FTAs, the rankings of equilibrium national welfare across club structures for $n = 4$ are as listed in Table B.6. In particular, welfare rankings are affected by the (a)symmetry of the structures and the densities of the clubs.*

[Insert Table B.5 here.]

[Insert Table B.6 here.]

As for $n = 3$, in the presence of FTAs, the formation of climate clubs will lead to higher carbon taxes in the member countries (environmental effect) and free trade among them (trade effect in the presence of FTAs).

In the absence of FTAs, as before, structures in which all countries are (in)directly connected with each other (i.e., structures V, VII, VIII, IX, X and XI) are equivalent in terms of policy decision making. These structures also preserve the symmetry among the countries since in every country, the carbon tax is chosen to maximize bilateral welfare and the trade tariff is chosen to maximize individual welfare. Moreover, the carbon tax in these structures is higher than that in a less cooperative structure (i.e., structures I and IV), while the tariff rate is lower. The overall effect is dominated by the benefits from lower trade barriers and so every country in a more cooperative structure is better off than in a less cooperative one, just as when $n = 3$.

When FTAs are part of the club treaties and the club structure is symmetric, we find that national welfare increases as the number of disjoint clubs/countries decreases (such as structures I, IV and IX), and stays the same for structures in which no two clubs are disjoint (such as structures IX and XI). Here, the benefits from free trade dominate the costs associated with higher carbon taxes when countries become (in)directly connected with each other. This is in line with the findings in Result 3.4 for $n = 3$.

When FTAs are part of the club treaties and the club structure is asymmetric, we find that it is no longer the case that the more climate clubs a country joins, the better off it is. More specifically, when the structure is hub-and-spoke or involves not many clubs that connect all countries (in)directly (such as structures V, VII and VIII), all countries bear the same level of carbon taxes. The benefits of free trading still dominate and countries joining more clubs enjoy higher welfare than others. On the other hand, in a structure with dense clubs (such as structure X), all countries also bear the same level of carbon taxes. But now the benefits of free trading with one additional country are outweighed by the losses in producer surplus and government revenues since every country is already free trading with most of the others. This makes countries joining less clubs be better off than the others.

With these welfare results, we derive the equilibrium structures according to different notions.

Result 3.7. *In the absence of FTAs, the equilibrium structures for $n = 4$ are as listed in Table B.7.*

Result 3.8. *In the presence of FTAs, the equilibrium structures for $n = 4$ are as listed in Table B.8.*

When club treaties do not involve FTAs, from Result 3.5, we know that all structures with interior solutions (i.e., structures I, IV, V, VII to XI) preserve the symmetry among the countries in the policy making stage. Therefore, in each structure, all countries have the same deviation incentive. Meanwhile, all structures in which the countries are (in)directly connected with each other (i.e., structures V, VII to XI) are equivalent in policy making, and any of them has at least two more clubs than structure IV, and structure IV has two more clubs than structure I. This implies that in the absence of FTAs, we could put the structures into three categories: category A - structures V, VII to XI (the most cooperative), category B - structure IV (the second most cooperative), and category C - structure I (the non-cooperative). Result 3.5 tells us that the more cooperative a structure is, the better off is each country. Considering all of these points, we see that when structures in categories A and B have interior solutions, the latter will neither be GS nor CPNE since three or more symmetric countries will want to deviate together to a structure in A; when structures in categories B and C have interior solutions, again the latter will neither be GS nor CPNE since countries in I will want to deviate together by forming two clubs. The finding is in line with that for $n = 3$.

When club treaties do involve FTAs, from Result 3.6, we see that just as in the case of $n = 3$, the geometric asymmetry plays a role in the equilibrium. When two symmetric structures I and XI have interior solutions, every country enjoys higher welfare in the connected one (structure XI) and so the non-cooperative one (structure I) is not GS/CPNE. When asymmetric structures have interior solutions, the asymmetry allows the country(ies) joining more clubs in a not relatively sparse structure (i.e., structure V, VII or VIII) to enjoy higher welfare than 1) others in the same structure and 2) being in a less asymmetric structure with either more or less clubs. In a dense structure X, the opposite holds. Therefore, for parameter ranges in which asymmetric structures that are not too dense have interior solutions, a structure is BS if no two countries joining less clubs find it profitable to deviate. For parameter ranges in which symmetric structures and the asymmetric dense structure have interior solutions, the latter is BS if the two countries joining more clubs do not find it profitable to deviate to a symmetric structure. The result is more general than that for $n = 3$ since there is no asymmetric dense structure with three countries.

3.5 Summary

In this section, we summarize the results from the analysis above.

3.5.1 Equilibrium Stability and Efficiency

We examine the equilibrium structures that are most stable and efficient, as well as the structures generated from different stability concepts. Tables B.9 and B.10 list the rankings of global welfare across structures.

[Insert Table B.9 here.]

[Insert Table B.10 here.]

Comparison between stability and efficiency

We see that the formation of climate clubs has two effects on national policies: 1) environmental effect: member countries bear higher carbon taxes than non-members; and 2) trade effect: member countries enjoy free trade with each other (when FTAs are part of the club treaties).

The subsequent welfare rankings are results of the overall effects on policies. More specifically, when club treaties do not involve FTAs, we see that: 1) all structures with interior solutions preserve the symmetry among the countries; 2) all structures in which the countries are (in)directly connected with each other via overlapping clubs generate the same policies; 3) structures with standalone countries do not have interior solutions; and 4) the more connected a structure is, the better off each country is. Therefore, structures in the absence of FTAs can be categorized based on the connectedness/cooperativeness of the countries: non-cooperative (such as structure I for $n = 3$ and 4, in which no clubs exist), partially-cooperative (such as structure IV for $n = 4$, in which all countries are involved in some clubs, but not all of them are (in)directly connected with each other via the clubs) and fully-cooperative (such as structures III and IV for $n = 3$ and structures V, VII to XI for $n = 4$, in which all countries are (in)directly connected with each other via overlapping clubs). It is then easy to see that a fully cooperative structure is not only immune to bilateral and multilateral deviations, but also maximizes global welfare.

Remark 3.1. *When club treaties do not involve FTAs, a fully cooperative structure is stable against all possible deviations and efficient.*

Now let's turn to clubs with FTAs. From Sections 3.3 and 3.4, we know that the national welfare and stability results would be dependent on the geometric symmetry and density of the structures. More specifically, among structures in which all countries are geometrically symmetric (such as structures I and IV for $n = 3$, and structures I, IV, IX and XI for $n = 4$), one could again categorize based on the connectedness/cooperativeness of the countries[7], and a country is better off in a more connected/cooperative structure. For structures in which countries are geometrically asymmetric, the density of the clubs plays a role. That is, in asymmetric structures that do not have a high density of clubs (such as structures II and III for $n = 3$, or structures V, VII and VIII for $n = 4$), a country joining more clubs would be better off than others joining less in the same structure, as well as those in a less asymmetric structure. For asymmetric structures with a high density of clubs (such as structure X for $n = 4$), a country joining less clubs would be better off than others joining more in the same structure, as well as those in a less asymmetric structure. On the other

[7]For example, when $n = 4$, the three categories are defined as before: non-cooperative (structure I), partially-cooperative (structure IV) and fully-cooperative (structures IX and XI).

hand, we find that the efficient structure in any parameter range would always be the one with the most clubs (see Tables B.9 and B.10). Therefore, whether or not stability would be in line with efficiency depends on which structures have interior solutions in the same parameter range.

Remark 3.2. *When club treaties involve FTAs,*

- *if only structures in which all countries are geometrically symmetric have interior solutions, then the most connected/cooperative structure would be stable against all possible deviations and efficient;*
- *if only structures with geometrically asymmetric countries and a not very high density of clubs have interior solutions, then the structure with the most clubs could be stable against all possible deviations and efficient;*
- *otherwise, a structure with less clubs may be stable against all possible deviations but not efficient.*

Comparisons between Different Stability Concepts

We now compare the equilibrium results based on different stability concepts. By definition, we know that $GS \subseteq CPNE \subseteq BS \subseteq NE$.

When club treaties do not involve FTAs, non-cooperative and partially/fully cooperative structures have interior solutions, in all of which the countries are geometrically symmetric. Meanwhile, all cooperative/connected structures are equivalent in terms of policy making. Therefore, we only need to consider the deviation between the non-cooperative and cooperative structures. Since the difference in the number of clubs between these two types of structures is at least two, then $BS = NE$. Moreover, since all countries have the same incentives to deviate from the non-cooperative structure to cooperative ones, then all profitable deviations are self-enforcing and $GS = CPNE \subset BS = NE$.

When club treaties include FTAs, we know that geometric asymmetry plays a role, and connected structures are no longer equivalent in terms of policy making and thus do not produce the same welfare results. If two structures with geometrically asymmetric countries have interior solutions and the difference in the number of clubs between them is just one, then $BS \subset NE$ is possible whenever a bilateral deviation is profitable. Moreover, we know that among structures that preserve the symmetry among the countries, the more clubs a structure has, the higher welfare each country enjoys. Therefore, it is profitable for groups of countries to deviate from the symmetric structure with less clubs to the one with more clubs. Now let's look at the parameter ranges in which an asymmetric structure and symmetric structures with both more and less clubs than the asymmetric structure have interior solutions. In these cases, some of the countries in the asymmetric structure enjoy higher welfare than those in the symmetric structure with more clubs, while others enjoy lower welfare. The opposite holds for the comparison between the asymmetric structure and the symmetric structure with less clubs. Thus, a group deviation from a symmetric structure with less clubs than the asymmetric structure to another symmetric structure with more clubs will not be self-enforcing since in the symmetric structure with more clubs, those enjoying lower welfare than in the asymmetric structure would like to further deviate together to the asymmetric structure. Hence, it is possible that $GS \subset CPNE$.

Remark 3.3. *The equilibrium results based on $CPNE$ and GS will be different when FTAs are included as part of the club treaties and geometric asymmetry makes the countries have different deviation incentives.*

3.5.2 Effects of the Formation of Multiple Clubs

As our model departs from the literature by considering multiple overlapping climate clubs, we want to know what effects such a consideration brings. We note that if at most one club is allowed to form, then only structure I or XI would be feasible.

We see that when club treaties involve FTAs, the formation of multiple clubs allows structures with geometrically asymmetric countries and not a very high density of clubs to have interior solutions, be stable and efficient for emission coefficients that are low. It also allows some asymmetric structures to be more stable though less efficient than the fully connected structure for emission coefficients that are intermediate. When club treaties do not involve FTAs and emission coefficients are intermediate, the formation of multiple clubs allows partially-cooperative structures to have interior solutions when fully-cooperative structures fail, which are also more stable and efficient than the non-cooperative structure. Overall, we have the following.

Remark 3.4. *The formation of multiple climate clubs improves stability of the equilibrium structures for emissions coefficients that are not large.*

3.5.3 Effects of FTAs

The results so far show that the presence of FTAs plays an important role in equilibrium outcomes. Here, we summarize the changes when the club treaties move from no-FTA to FTA-included.

We see that no matter multiple clubs are allowed or not, the inclusion of FTAs enables the fully connected structure in which every pair of countries forms a club with each other to have interior solutions, be stable against all possible deviations and efficient whenever emissions coefficients are sufficiently large. Note that this structure is equivalent to a grand coalition structure in which a multilateral club involves all countries. Its emergence enlarges the stable and efficient region of the equilibrium. When multiple clubs are allowed to form, the presence of FTAs would enable structures with geometrically asymmetric countries, which are less connected/cooperative than the grand coalition, to emerge in equilibrium. In particular, these structures are stable against all deviations and efficient for even lower values of emissions coefficients than in the one-club case. For some intermediate values of emissions coefficients, these structures are more stable though may be less efficient than the grand coalition, which is the equilibrium structure without FTAs.

Remark 3.5. *Regardless of the consideration of multiple clubs, the presence of FTAs in club treaties allows the grand coalition to be stable against all possible deviations and efficient for all emissions coefficients sufficiently large, which enlarges the stable and efficient region of the equilibrium. When multiple clubs are considered, the emergence of structures with geometrically asymmetric countries in the equilibrium further enlarges this region to include lower values of emissions coefficients and improves the stability of the equilibrium though not the efficiency for some intermediate values of emissions coefficients.*

CHAPTER 4

STRATEGIC EFFECTS OF FUTURE ENVIRONMENTAL POLICY COMMITMENTS: CLIMATE CHANGE, SOLAR RADIATION MANAGEMENT AND CORRELATED AIR POLLUTANTS

4.1 Introduction

If in the near future multiple instruments are available (e.g., carbon pricing and solar radiation management) and governments are able to commit to a particular type of policy instrument, which instrument will they prefer? Are there clear game theoretic predictions of play? This paper is an attempt to answer these questions.

There is not much reason for optimism with respect to the prospects of implementation of an effective, cooperative, international agreement to curb the evils produced by climate change. The Kyoto protocol has not produced enthusiastic results and a post-Kyoto agreement does not promise to be much different. The high national costs associated with mitigation of greenhouse gas emissions appear to be the main culprit.

Revealed preference informs us that some nations prefer the status quo of no significant mitigation of greenhouse gas emissions to a commitment to reduce greenhouse emissions by a significant percentage amount relative to 1990 levels. However, this fact does not rule out the possibility that governments, which have rejected Kyoto as well as those that may reject a post-Kyoto agreement, are currently contemplating adopting cheaper alternatives to mitigation of greenhouse gas emissions in order to reduce their potential damages caused by climate change. In fact, there appears to be credence in the scientific community that some nations are seriously considering producing climate change engineering products – such as solar radiation management (SRM) generated by injections of sulfate aerosols into the stratosphere – that may effectively control the global temperature (see, e.g., the discussion of the scientific findings in [42]). One must also account for some of the potential negative effects associated with climate change engineering. SRM, for example, is expected to produce droughts, ozone depletion and to change the color of our blue skies.

This paper studies some of the effects associated with uncoordinated policy commitments with respect to provision of SRM) relative to policy commitments for mitigation of greenhouse gas (e.g., carbon dioxide) emissions. We envision environmental policy making in a future global economy in which SRM is a proven and mature technology, which can be deployed at will and unilaterally by any nation. Among other things, we analyze whether there will be an incentive for a nation to be a policy leader in mitigation of carbon dioxide emissions or SRM provision, where the nations are perfectly informed about the benefits and costs of providing SRM. Solar radiation management is expected to produce droughts, ozone depletion and to change the color of our blue skies.

The fact that SRM may soon prove to be a cheaper and effective alternative to mitigation of carbon dioxide emissions implies that unilateral action in SRM will not only be credible (see, e.g., [8]), but also that nations may then wish to commit to SRM policies and subject their carbon mitigation policies to SRM policy commitments. Recent and noteworthy contributions to the literature have considered some of the potential reactions we may observe with future implementation of geoengineering technologies (e.g., [33], [41], [42], [43], [44], [50]). Moreno-Cruz (2011) examines non-cooperative games in which two nations are either symmetric or asymmetric with respect to drought damages. In the symmetric game, he finds that the prospect of SRM will create greater incentives for free riding on carbon mitigation. When nations are asymmetric, he finds that SRM provision can induce inefficiently high levels of mitigation. Millard-Ball studies the impact of geoengineering deployment on the formation of a mitigation agreement. He shows that a credible unilateral threat of utilizing geoengineering may strengthen global abatement and lead to a self-enforcing climate treaty with full participation. Urpelainen shows that geoengineering may induce significant reductions in emissions in the present if it produces severe negative externalities, since the latter may lead to a very harmful geoengineering race in the future. If the externalities are not overly severe, unrestricted utilization of geoengineering can be globally beneficial. Our paper contributes to this literature in at least three significant ways. First and foremost, we examine the effects associated with strategic environmental policy commitments, whereby SRM policy may precede carbon policy. This may indeed occur in the future when SRM technology is mature. SRM policy may be (politically or even socially) cheaper and easier to implement than carbon policy. Our motivation here is therefore to consider a likely future event and then make a prediction concerning the equilibrium policies. As in the papers cited above, we assume that SRM provision generates global damages - in our setting SRM produces drought damages and the drought damage function is increasing at an increasing rate.

Second, our model accounts for the fact that emissions of carbon dioxide are correlated with emissions of sulfur dioxide due to important common sources, such as energy production. Our model builds on [16].[1] As in [16], sulfur dioxide emissions cause acid rain damage in the emitting nation. We show that the instruments a nation utilizes to control carbon and sulfur dioxide emissions are strategic complements. Hence, whenever SRM provision leads to an increase in carbon emissions, it also leads to an increase in sulfur emissions, with a resulting increase in acid rain damage. Finally, unlike the cited papers, we examine environmental policy making within a general equilibrium framework. This will enable us to see how consumers and industry emitters respond to strategic policy choices made by the governments.

[1]Our framework can also be seen as an extension of the impure public good model advanced by [20] to a context in which there are two impure public goods, namely, SRM provision and mitigation of carbon dioxide emissions. SRM provision yields global pure public good benefits, but entails national-specific drought damages. Reduction of carbon emissions also yields global pure public good benefits, but entails national-specific costs in terms of reduction of the consumer surplus associated with energy consumption. Our analysis makes a contribution to the public goods literature in that we consider both simultaneous and sequential strategic interactions between these two types of impure public goods.

4.2 Modeling Strategies and Brief Discussion of Main Results

We consider a global economy consisting of two nations, which are identical in all respects, except for the drought and acid rain damage functions. This is a modeling strategy. We wish to highlight the effects that differences in both drought and acid rain damages may promote in the formulation of non-cooperative carbon and SRM environmental policies and on the incentives for policy commitments.

Each nation has three policy instruments at its disposal; namely, SRM provision and carbon and sulfur pollution permits. Our choice of pollution permits as the means to price emissions is motivated by the Kyoto Protocol, the European Union Emissions Trading System and the 1990 US Clean Air Act Amendments, which created a national program in tradable sulfur dioxide emission permits.

Although we consider the making of uncoordinated environmental policies in a future time when policy makers have SRM at their disposal, our analysis involves a single period. The various timings of the games examined in this paper are strictly motivated by individual costs and benefits of policy commitments. We wish to predict which timing is likely to emerge in equilibrium. The timings are not motivated by the historical evolution of the utilization of environmental policy instruments. An alternative and interesting avenue for research is to explicitly consider an intertemporal model in which the sequence of policy instruments mimics the historical evolution of environmental policy, with sulfur pollution permits preceding carbon pollution permits and the latter preceding SRM. In such a case, the sequencing is exogenous and one considers the impacts associated with the sequential introduction of environmental policy instruments.

We start our analysis by considering two benchmarks in which environmental policy making with respect to SRM provision and carbon-permit quotas are chosen simultaneously: (i) uncoordinated policy making; and (ii) fully coordinated policy making. Following [16], in all games examined in this paper, we assume that environmental policy with respect to sulfur quotas are chosen after the other two types of environmental policy instruments. One can explain this on the basis that our analysis concerns likely future events when climate change policies receive priority in environmental policy making. This is a global setting in which there is global consensus that climate change dangers need to be addressed and in which nations are endowed with various instruments to control the negative effects caused by climate change. The main purpose of this paper is to examine strategic effects for the mostly likely environmental policy scenarios in such a futuristic world.

In the next two games, we examine sequential choices of SRM and carbon quotas, but the nations make simultaneous choices of SRM provision or carbon pollution quotas. These games consist of three stages. In game III, the nations make choices with respect to carbon quotas in the first stage, make choices with respect to SRM provision in the second stage and make choices with respect to sulfur quotas in the last stage. In game IV, the first stage consists of simultaneous SRM choices and the second stage of simultaneous choices of carbon quotas.

The last two games consider the effects of policy leadership in carbon or SRM policy. These are four stage games. In game V, a nation chooses its carbon quota in the first stage, the other nation chooses its carbon quota in the second stage, and the two nations simultaneously choose SRM provision in the third stage and simultaneously choose sulfur quotas in the last stage. In game VI, a nation chooses its SRM provision in the first stage, the other nation chooses its SRM provision in the second stage, and the two nations simultaneously choose carbon quotas in the third stage and sulfur quotas in the last stage.

We have several important findings. We show that if both nations simultaneously commit to carbon permit policies, national SRM levels rise with the carbon quotas. The global temperature rises following each unit increase in the global carbon quota. If, on the other hand, both nations simultaneously commit to SRM policies, national carbon quotas rise with national SRM levels. The global temperature falls following each unit increase in the global SRM level. We also find that a nation always has an incentive to be a policy leader in either carbon or SRM policy. For various values of parameters of utility and technological functions, we can also show that a nation prefers being a leader in carbon policy to being a leader in SRM policy, but the globe prefers leadership in SRM policy to leadership in carbon policy. In addition, the globe prefers simultaneous policy commitments in SRM policy to simultaneous policy commitments in carbon policy. As for SRM provision, we show that it is overprovided whenever carbon policy is determined prior to SRM policy and it is underprovided otherwise. Global carbon emissions are always larger than the globally efficient amount, but the second lowest level of global carbon emissions is observed in the setting in which there is leadership in SRM policy. Carbon emissions follow the same pattern as sulfur emissions.

The remainder of the chapter is organized as follows. Section 4.3 describes the general equilibrium framework, Section 4.4 provides the analysis for the various policy games, Section 4.5 considers whether or not the settings involve first-mover advantage, Section 4.6 offers key results of comparisons across equilibria, and Section 4.7 concludes.

4.3 General Equilibrium Framework

Consider a global economy that, for simplicity, consists of two nations, indexed by $j, j = 1, 2$. We normalize the population of each nation, letting it be equal to 1. We assume that both nations suffer from droughts caused by solar radiation management (SRM), global warming caused by emissions of carbon dioxide and acid rain caused by emissions of sulfur dioxide. Carbon dioxide and sulfur dioxide emissions are by-products of energy production in each nation. Solar radiation management, through injections of sulfur aerosols, is provided by each national government.

Let $H^T(C, M)$ denote the harm function associated with global warming in each nation, where $C = \sum_{j=1}^{2} C_j$ and $M = \sum_{j=1}^{2} M_j$ are global levels of carbon dioxide and SRM, respectively. For concreteness, we shall assume that $H^T(C, M) = h^T(C - M)^2$, where $h^T > 0$. Let $H_j^A(\theta_j, S_j)$ be the harm function of acid rain deposition in nation j, where

S_j is the level of sulfur dioxide in nation j and $1 \geq \theta_j > 0, j = 1, 2$ denote its sensitivity to acid rain damage. We assume that $H_j^A(\theta_j, S_j) = h^A \theta_j S_j^2$, where $1 \geq h^A > 0$. The drought damage caused by SRM in nation j is represented by $H_j^D(\delta_j, M)$, where $1 \geq \delta_1 \geq \delta_2 > 0, j = 1, 2$ denotes its sensitivity to drought damage. We assume that $H_j^D(\delta_j, M) = \delta_j h^D M^2$, where $1 \geq h^D > 0$.[2] Nation j's total level of environmental damages is

$$H_j = H_j(C, M, S_j) = H^T(C, M) + H_j^A(\theta_j, S_j) + H_j^D(\delta_j, M). \tag{4.1}$$

The representative consumer in nation j consumes x_j units of composite good (numeraire), e_j units of energy and is harmed by H_j units of environmental damages. Let $u_j(x_j, e_j, H_j)$ denote consumer j's utility function. Assume that u_j is quasi-linear (linear in x_j), and increasing and strictly concave in e_j. More precisely, we assume that $u_j(x_j, e_j, H_j) = x_j + f(e_j) - H_j$ where $f(e_j) = e_j(b - ae_j)$ and $b \geq 1 \geq a > 0$.[3] The representative consumer?s income is denoted w_j, where

$$w_j = \bar{x}_j + p_{C_j} Q_{C_j} + \pi_j - K^M(M_j). \tag{4.2}$$

The consumer is initially endowed with $\bar{x}_j$ units of the numeraire good, is the sole shareholder of profits earned by her nation?s energy industry, π_j, is the recipient of the revenues generated with sale of pollution permits, $p_{C_j} Q_{C_j} + p_{S_j} Q_{S_j}$ (p_{C_j} and p_{S_j} are prices of carbon and sulfur permits and Q_{C_j} and Q_{S_j} are the carbon and sulfur quotas in nation j) and pays a tax equal to her nation's cost of provision of SRM, $K^M(M_j)$. We shall assume

[2] Quadratic pollution damage functions are widely used in the environmental economics literature that considers game-theoretic applications. Provided the damage functions are strictly convex, the nations' policies will be strategic. The results with different strictly convex specifications of the damage functions will be qualitatively identical to the ones we obtain in this paper.

[3] The quasi-linear utility function characterization is frequently used in general-equilibrium models because of its desirable aggregation properties (demand side). In addition, quasi-linearity together with the assumption that the utility function from energy consumption is quadratic yields a linear demand function for energy. This type of demand function is commonly used in the regulation literature. The assumption that the utility function is separable in energy consumption and pollution damages is also standard in the environmental economics literature.

that $K^M(M_j) = k^M M_j^2$, where $1 \geq k^M > 0$. For future use, it is important to note that $C = Q_{C_1} + Q_{C_2} \equiv Q_C$ and $S_j = Q_{S_j}, j = 1, 2$.

Throughout, we assume that markets are competitive, with consumers and energy firms taking prices as given. Consumers also take their incomes and the levels of pollution damages as given. Consumer j chooses non-negative $\{x_j, e_j\}$ to maximize $u_j(x_j, e_j, H_j)$ subject to $x_j + p_{e_j} e_j = w_j$, where p_{e_j} is nation j's energy price. This is equivalent to choosing $e_j \geq 0$ to maximize $f(e_j) - p_{e_j} e_j$. Assuming interior solutions, the first-order condition is:

$$f'(e_j) - p_{e_j} = 0. \tag{4.3}$$

Condition (3) informs us that the optimal level of energy to be consumed is the one at which the marginal utility from energy equates the price of energy. Since the second-order condition is $f"(e_j) = -2a < 0$, the solution to the consumer's maximization problem is unique. Equation (3) yields $e_j(p_{e_j}) = (b - p_{e_j})/(2a)$, consumer j's energy demand. Her demand for the numeraire good is $x_j(p_{e_j}, w_j) = w_j - p_{e_j} e_j(p_{e_j})$ and $v_j(p_{e_j}, w_j, H_j) = x_j(p_{e_j}, w_j) + f(e_j(p_{e_j})) - H_j$ is her indirect utility function.

In nation j, the energy industry's profit function is

$$\begin{aligned}
&\pi_j(E_j, R_{C_j}, R_{S_j}) \\
&= p_{e_j} E_j - p_{C_j}(E_j - R_{C_j}) - p_{S_j}(E_j - R_{S_j}) - K^E(E_j) - K^C(R_{C_j}) - K^S(R_{S_j}) \\
&= p_{e_j} E_j - p_{C_j} \Delta_j^C(E_j, R_{C_j}) - p_{S_j} \Delta_j^S(E_j, R_{S_j}) - K^E(E_j) - K^C(R_{C_j}) - K^S(R_{S_j})
\end{aligned}$$

where E_j, R_{C_j} and R_{S_j} are levels of energy production, carbon dioxide reduction and sulfur reduction, respectively. We denote by $\Delta_j^C(E_j, R_{C_j}) \equiv E_j - R_{C_j}$ and $\Delta_j^S(E_j, R_{S_j}) \equiv E_j - R_{S_j}$ the the quantities of carbon and sulfur permits demanded by nation j's energy industry, respectively. To simplify notation, we shall assume that the size of the energy industry in each nation is equal to 1 and thus refer to it as an energy ?firm?. We also assume that the costs of energy production, carbon dioxide reduction and sulfur dioxide reduction are $K^E(E_j) = k^E E_j^2$, $K^C(R_{C_j}) = k^C R_{C_j}^2$ and $K^S(R_{S_j}) = k^S R_{S_j}^2$, where $1 \geq k^E > 0$, $1 \geq k^C > 0$ and $1 \geq k^S > 0$.[4] Each firm chooses non-negative $\{E_j, R_{C_j}, R_{S_j}\}$ to maximize π_j, taking all prices as given. The first-order conditions for interior solutions are

$$p_{e_j} - p_{C_j} - p_{S_j} - \frac{dK^E}{dE_j} = 0, \tag{4.4a}$$

$$p_{C_j} - \frac{dK^C}{dR_{C_j}} = 0, \tag{4.4b}$$

$$p_{S_j} - \frac{dK^S}{dR_{S_j}} = 0. \tag{4.4c}$$

Condition (4a) informs us that the optimal amount of energy to be produced in a nation should equate the marginal revenue to the sum of marginal production and regulatory costs of energy production. Equation (4b) states that the optimal level of carbon abatement should equate the marginal revenue from carbon abatement (i.e., the marginal cost saving in expenditure on carbon permits) to the marginal cost of carbon abatement. Equation (4c) is similar; it equates marginal revenue from sulfur abatement to the marginal cost of sulfur abatement.

Solving the system of equations (4a) to (4c), we obtain nation j's energy supply function, $E_j(p_{e_j}, p_{C_j}, p_{S_j}) = (p_{e_j} - p_{C_j} - p_{S_j})/2k^E$, and the carbon and sulfur abatement supply functions, $R_{C_j}(p_{C_j}) = p_{C_j}/2k^C$ and $R_{S_j}(p_{S_j}) = p_{S_j}/2k^S$, respectively. Thus, we have (for $j = 1, 2$)

[4] The cost functions are assumed to be quadratic for tractability purposes and in order to generate linear supply functions. Provided the cost functions are increasing and strictly convex and the marginal willingness to pay for the energy good is sufficiently high, an interior and unique equilibrium is guaranteed. The results under other strictly convex specifications of the costs functions will be qualitatively identical to the ones we obtain in the text.

$$\frac{\partial E_j}{\partial p_{e_j}} = -\frac{\partial E_j}{\partial p_{C_j}} = -\frac{\partial E_j}{\partial p_{S_j}} = \frac{1}{2k^E} > 0, \tag{4.5a}$$

$$\frac{dR_{C_j}}{dp_{C_j}} = \frac{1}{2k^C} > 0, \tag{4.5b}$$

$$\frac{dR_{S_j}}{dp_{S_j}} = \frac{1}{2k^S} > 0. \tag{4.5c}$$

As expected, conditions (5a) inform us that each energy firm?s supply function is increasing the price of energy and decreasing in the prices of carbon and sulfur permits. Conditions (5b) reveal that each firm supplies more carbon abatement as the price of the carbon permit in its nation increases. Conditions (5c) are similar to conditions (5b): they state that each firm produces more sulfur abatement as the price of its nation's sulfur permit increases. Combining the definitions of the quantities of carbon and sulfur permits demanded with conditions (5a) - (5c) yields

$$\frac{\partial \Delta_j^C}{\partial p_{e_j}} = \frac{\partial \Delta_j^S}{\partial p_{e_j}} = -\frac{\partial \Delta_j^C}{\partial p_{S_j}} = -\frac{\partial \Delta_j^S}{\partial p_{C_j}} = \frac{1}{2k^E} > 0, \tag{4.6a}$$

$$\frac{\partial \Delta_j^C}{\partial p_{C_j}} = -\frac{1}{2k^E} - \frac{1}{2k^C} < 0, \tag{4.6b}$$

$$\frac{\partial \Delta_j^S}{\partial p_{S_j}} = -\frac{1}{2k^E} - \frac{1}{2k^S} < 0. \tag{4.6c}$$

Conditions (6a) inform us that the quantities of carbon and sulfur permits demanded rise with the price of energy. This is natural since the firms will expand energy production as the price of energy increases. Conditions (6a) also reveal the quantities of carbon and sulfur permits demanded are complements, since the quantity of carbon permits demanded by firm j falls as the price of nation j's sulfur permit rises and the quantity of sulfur permits demanded by firm j falls as the price of nation j's carbon permit rises. Conditions (6b) and (6c) state that the demands for carbon and sulfur permits fall as their respective prices increase.

Market-clearing conditions for the national energy market, and carbon and sulfur permit markets, respectively, are as follows:

$$e_j(p_{e_j}) = E_j(p_{e_j}, p_{C_j}, p_{S_j}), \tag{4.7a}$$

$$\Delta_j^C(p_{e_j}, p_{C_j}, p_{S_j}) = Q_{C_j}, \tag{4.7b}$$

$$\Delta_j^S(p_{e_j}, p_{C_j}, p_{S_j}) = Q_{S_j}. \tag{4.7c}$$

Conditions (7a) inform us that in each nation the demand for energy must be equal to the supply of energy. Conditions (7b) state that in each nation, the demand for carbon permits must be equal to the supply of carbon permits. Conditions (7c) are similar to conditions (7b): in each nation, the demand for sulfur permits equals the supply of sulfur permits. Solving the system of equations (7a) - (7c), we obtain the price functions in terms of pollution quotas:

$$p_{e_j}(Q_{C_j}, Q_{S_j})) = \frac{b(k^C + k^E + k^S) - 2b(k^C Q_{C_j} + k^S Q_{S_j})}{a + k^E + k^C + k^S}, \tag{4.8a}$$

$$p_{C_j}(Q_{C_j}, Q_{S_j})) = \frac{bk^C - 2k^C[(a + k^E + k^S)Q_{C_j} - k^S Q_{S_j}]}{a + k^E + k^C + k^S}, \tag{4.8b}$$

$$p_{S_j}(Q_{C_j}, Q_{S_j})) = \frac{bk^S + 2k^S[k^C Q_{C_j} - (a + k^C + k^E)Q_{S_j}]}{a + k^E + k^C + k^S}. \tag{4.8c}$$

Then the following comparative static results are immediate:

$$\frac{\partial p_{e_j}}{\partial Q_{C_j}} = -\frac{2bk^C}{a+k^E+k^C+k^S} < 0, \tag{4.9a}$$
$$\frac{\partial p_{C_j}}{\partial Q_{C_j}} = -\frac{2k^C(a+k^E+k^C)}{a+k^E+k^C+k^S} < 0, \tag{4.9b}$$
$$\frac{\partial p_{S_j}}{\partial Q_{C_j}} = \frac{2k^Ck^S}{a+k^E+k^C+k^S} > 0, \tag{4.9c}$$
$$\frac{\partial p_{e_j}}{\partial Q_{S_j}} = -\frac{2bk^S}{a+k^E+k^C+k^S} < 0, \tag{4.9d}$$
$$\frac{\partial p_{C_j}}{\partial Q_{S_j}} = \frac{2k^Ck^S}{a+k^E+k^C+k^S} > 0, \tag{4.9e}$$
$$\frac{\partial p_{S_j}}{\partial Q_{S_j}} = -\frac{2k^S(a+k^E+k^C)}{a+k^E+k^C+k^S} < 0. \tag{4.9f}$$

We summarize our findings about the equilibrium price functions in the following proposition.

Proposition 4.1. *$\frac{\partial p_{e_j}}{\partial Q_{C_j}} < 0$, $\frac{\partial p_{C_j}}{\partial Q_{C_j}} < 0$, $\frac{\partial p_{S_j}}{\partial Q_{C_j}} > 0$, $\frac{\partial p_{e_j}}{\partial Q_{S_j}} < 0$, $\frac{\partial p_{C_j}}{\partial Q_{S_j}} > 0$ and $\frac{\partial p_{S_j}}{\partial Q_{S_j}} < 0$, $j = 1, 2$.*

We are now ready to write consumer j's indirect utility as function of pollution quotas and SRM provision levels:

$$\begin{aligned} & v_j(Q_{C_j}, Q_{C_{-j}}, Q_{S_j}, M_j, M_{-j}) \\ & = \bar{x}_j + f_j(e_j(p_{e_j}(\cdot))) - K^E(E_j(p_{e_j}(\cdot)), p_{C_j}(\cdot)), p_{S_j}(\cdot))) - K^C(R_{C_j}(p_{C_j}(\cdot))) \\ & - K^S(R_{S_j}(p_{S_j}(\cdot))) - K^M(M_j) - H_j^A(\theta_j, Q_{S_j}) - H^T(Q_C, M) - H_j^D(\delta_j, M) \end{aligned} \tag{4.10}$$

where $p_{e_j}(\cdot) = p_{e_j}(Q_{C_j}, Q_{S_j})$, $p_{C_j}(\cdot) = p_{C_j}(Q_{C_j}, Q_{S_j})$, $p_{S_j}(\cdot) = p_{S_j}(Q_{C_j}, Q_{S_j})$ and $M = M_j + M_{-j}, j = 1, 2$. We let $-j = 1$ if $j = 2$ and $-j = 2$ if $j = 1$.

4.4 Effects of Environmental Policy Commitments

We now examine the effects of timing in environmental policy making. We consider six sequential policy games as described below. In all games, we assume that the national governments select their sulfur quotas at the last stage. The equilibrium concept utilized is subgame perfection. Table 4.1 summarizes the timings of the six games considered in the analysis.

We establish two benchmarks under which policy making on carbon quotas and SRM levels occur simultaneously in the first stage of two-stage games: (1) uncoordinated policy making; and (2) coordinated policy making. Sulfur quotas are chosen simultaneously in the second stage of the game. These benchmarks will enable us to capture both the effects caused by the timing of policy making and the effects caused by non-cooperative behavior (i.e., departures from socially efficient behavior). The next two sequential games involve simultaneous choices of either (3) pollution quotas; or (4) SRM levels; in the first stage, with simultaneous choices of either SRM (in game (3)) or pollution quotas (in game (4)) in the second stage. Thus, unlike the games in which pollution quotas and SRM are chosen simultaneously, games (3) and (4) involve three stages, since sulfur quotas are simultaneously chosen in the third stage of the game. The last two sequential games involve four stages. In game (5), nation 1 chooses its carbon quota in the first stage, nation 2 chooses its carbon quota in the second stage, SRM levels are chosen simultaneously in the third stage and sulfur quotas are chosen simultaneously in the fourth stage. In game (6), nation 1 chooses its SRM level in the first stage, nation 2 chooses its SRM level in the second stage, carbon quotas are chosen simultaneously in the third stage and sulfur quotas are chosen simultaneously in the fourth stage.

Consider the last stage of any game, namely, the stage in which the nations choose sulfur quotas simultaneously after having observed the other policy choices, $\{Q_{C_j}, Q_{C_{-j}}, M_j, M_{-j}\}$. The optimization problem faced by nation j's government is to choose non-negative Q_{S_j} to

Table 4.1 – Environmental policy games.

Timing	Descriptions
Simultaneous SRM and Carbon Quotas	The four choices, both nations' carbon quotas and SRM levels, are made simultaneously.
1) Uncoordinated	Carbon quotas and SRM levels are coordinated between the two nations, respectively.
2) Coordinated	Carbon quotas and SRM levels are uncoordinated between the two nations, respectively.
Sequential SRM and Carbon Quotas	
No National Leadership	
3) First Stage: Carbon Quotas	Nations decide carbon quotas simultaneously in the first stage and choose SRM levels simultaneously in the second stage.
4) First Stage: SRM Levels	Nations choose SRM levels simultaneously in the first stage and decide carbon quotas simultaneously in the second stage.
National Leadership	
5) First Stage: Nation 1's Carbon Quota	In the first stage, nation 1 decides its carbon quota, followed by nation 2's carbon decision in the second stage. Both nations choose SRM levels simultaneously in the third stage.
6) First Stage: Nation 1's SRM Level	In the first stage, nation 1 chooses its SRM level, followed by nation 2's SRM choice in the second stage. Both nations decide carbon quotas simultaneously in the third stage.

maximize the national indirect utility function (9), taking $Q_{S_{-j}}$ as given. Assuming interior solutions, we obtain the following first-order conditions:

$$p_{S_j} - \frac{\partial H_j^A}{\partial Q_{S_j}} = 0, j = 1, 2. \tag{4.11}$$

Conditions (11) inform us that in each nation the amount of sulfur quota should be set at the level that equates the national permit price to the national marginal damage caused by acid rain. Since $H_j^A = h^A \theta_j Q_{S_j}^2$, we have

$$Q_{S_j} = p_{S_j}/2h^A\theta_j, j = 1, 2. \tag{4.12}$$

Substituting equations (12) into (8a)-(8c) yields

$$Q_{S_j} = \frac{k^S(b + 2k^C Q_{C_j})}{2[k^S(a + k^E + k^C) + h^A\theta_j(a + k^E + k^C + k^S)]}, \tag{4.13a}$$
$$p_{e_j} = \frac{b[k_S(k^E + k^C) + h^A\theta_j(k^E + k^C + k^S)] - 2ak^C(k^S + h^A\theta_j)Q_{C_j}}{k^S(a + k^E + k^C) + h^A\theta_j(a + k^E + k^C + k^S)}, \tag{4.13b}$$
$$p_{C_j} = \frac{k^C\{b(k^S + h^A\theta_j) - 2[k^S(a + k^E) + h^A\theta_j(a + k^E + k^S)]Q_{C_j}\}}{k^S(a + k^E + k^C) + h^A\theta_j(a + k^E + k^C + k^S)}, \tag{4.13c}$$
$$p_{S_j} = \frac{k^S h^A\theta_j(b + 2k^C Q_{C_j})}{k^S(a + k^E + k^C) + h^A\theta_j(a + k^E + k^C + k^S)}. \tag{4.13d}$$

Hence, we can now obtain comparative static results when sulfur quotas are chosen optimally by the national governments:

$$\frac{\partial p_{e_j}}{\partial Q_{C_j}} = -\frac{2ak^C(k^S + h^A\theta_j)}{k^S(a + k^E + k^C) + h^A\theta_j(a + k^E + k^C + k^S)} < 0, \tag{4.14a}$$
$$\frac{\partial p_{C_j}}{\partial Q_{C_j}} = -\frac{2k^C k^S(a + k^E) + 2h^A k^C\theta_j(a + k^E + k^S)}{k^S(a + k^E + k^C) + h^A\theta_j(a + k^E + k^C + k^S)} < 0, \tag{4.14b}$$
$$\frac{\partial p_{S_j}}{\partial Q_{C_j}} = \frac{2k^C k^S h^A\theta_j}{k^S(a + k^E + k^C) + h^A\theta_j(a + k^E + k^C + k^S)} > 0, \tag{4.14c}$$
$$\frac{\partial Q_{S_j}}{\partial Q_{C_j}} = \frac{k^C k^S}{k^S(a + k^E + k^C) + h^A\theta_j(a + k^E + k^C + k^S)} > 0. \tag{4.14d}$$

Unlike in Proposition 4.1, the sulfur quotas are now adjusted optimally when the carbon quotas are increased. Condition (14d) informs us that a nation's sulfur and carbon quotas are strategic complements. As a nation's carbon quota increases, its sulfur quota is also increased, implying that the inelastic curve representing the supply of sulfur permits shifts to the right when the carbon quota is increased. Since, as we have discussed after Proposition 4.1, the demand for sulfur permits shifts out when the carbon quota is increased, the final effect of increasing the carbon quota on the sulfur permit price is potentially ambiguous. However, as condition (14c) demonstrates, the net effect on the sulfur permit price is positive. The optimal adjustment in the sulfur quota reduces but does not eliminate the hike in the sulfur permit price relative to a situation where sulfur quota is not adjusted at all (i.e., when the sulfur quota policy is exogenous). The net effects of an increase in the carbon quota on the energy and carbon permit prices captured by conditions (14a) and (14b), respectively, are the expected ones. We summarize results (14a) - (14d) in the following proposition.

Proposition 4.2. $\frac{\partial p_{e_j}}{\partial Q_{C_j}} < 0$, $\frac{\partial p_{C_j}}{\partial Q_{C_j}} < 0$, $\frac{\partial p_{S_j}}{\partial Q_{C_j}} > 0$, $\frac{\partial Q_{S_j}}{\partial Q_{C_j}} > 0$, $j = 1, 2$.

4.4.1 Game I: Simultaneous Choices of Carbon Quotas and SRM Levels

In the first stage, the government of nation j chooses non-negative $\{Q_{C_j}, M_j\}$ to maximize (10) subject to $Q_{S_j} = Q_{S_j}(Q_{C_j}; \theta_j)$, taking $\{Q_{C_{-j}}, M_{-j}\}$ as given. Assuming interior solutions, the set of first-order conditions are (for $j = 1, 2$):

$$p_{C_j} = \frac{\partial H^T}{\partial Q_C}, \tag{4.15a}$$

$$\frac{dH^T}{dM} = \frac{dK^M}{dM_j} + \frac{dH_j^D}{dM} = 0. \tag{4.15b}$$

Conditions (15a) reveal that the carbon quota in nation j should be set at the level that equates the national carbon permit price to the national marginal damage of global warming. Conditions (15b) are similar in spirit, since the level of SRM that should be provided at nation j is the one under which the national marginal benefit from SRM provision is equal to the national marginal cost. The latter is the sum of the marginal cost of provision and the marginal damage caused by droughts in nation j.

Since $H^T = h^T(C - M)^2$, $H_j^D = \delta_j h^D M^2$ and $K^M = k^M M_j^2$, equations (15b) yield:

$$M_j = \frac{h^T Q_C[k^M + h^D(\delta_{-j} - \delta_j)]}{k^M(2h^D + 2h^T + k^M)} \tag{4.16}$$

where $-j = 2$ if $j = 1$ and vice versa. Equations (16) make it clear that the assumption that the equilibrium is interior requires us to consider situations where the national sensitivity indexes associated with the damage functions from droughts are not too different from each other. Henceforth, we shall assume that $\delta_2 + (k^M/h^D) > \delta_1 \geq \delta_2$. Furthermore, given our assumptions, the sufficient second order condition for this game is satisfied. The equilibrium is unique.[5]

4.4.2 Game II: Simultaneous Coordinated Choices

Suppose now that there is full coordination between the two nations regarding the choices of carbon quotas and SRM levels in the first stage of the game. Assume that a utilitarian bi-national environmental agency chooses non-negative $\{Q_{C_1}, Q_{C_2}, M_1, M_2\}$ to maximize global indirect utility $V(\cdot) = v_1(\cdot) + v_2(\cdot)$ (where $v_1(\cdot)$ and $v_2(\cdot)$ correspond to function (10) by setting $j = 1, 2$, respectively) subject to (13a). Assuming interior solutions, the first-order conditions are as follows ($j = 1, 2$):

$$p_{C_j} = 2(\frac{\partial H^T}{\partial Q_C}), \tag{4.17a}$$

$$-2(\frac{\partial H^T}{\partial M}) = \frac{dK^M}{dM_j} + \sum_{j=1}^{2} \frac{dH_j^D}{dM}. \tag{4.17b}$$

[5] The sufficient second order conditions are satisfied in all games examined in this paper. Hence, the equilibrium for each game is unique. These results are available from the authors upon request.

Condition (17a) tells us that the amount of carbon quotas in nation j is chosen to equate the national carbon permit price to the social marginal damage of global warming - that is, the sum of the national marginal damages from global warming. Condition (17b) state that nation j should provide SRM at a level that equalizes the global marginal benefit from SRM provision (i.e., the sum of national marginal benefits from a reduction in the global temperature) to the global marginal cost of SRM provision (i.e., the sum of national marginal production cost and the sum of national marginal damages from droughts caused by SRM provision).

4.4.3 Game III: Simultaneous Commitments on Carbon Permit Policies

Having observed $\{Q_{C_j}, Q_{C_{-j}}\}$, government j chooses non-negative M_j to maximize its indirect utility function (9), subject to (12a), taking M_{-j} as given in the second stage. Assuming interior solutions, the first-order conditions yield equations (15). These best-response functions reveal that SRM provision in nation j expands with the global quantity of carbon dioxide. Adding the best-response functions, one gets the global SRM level as function of the global carbon dioxide quantity. The global SRM quantity rises with the global carbon dioxide quantity, but at a rate that is less than one.[6]

In the first stage, the first-order conditions are ($-j = 2$ if $j = 1$ and vice versa):

$$p_{C_j} = \frac{\partial H^T}{\partial Q_C} + (\frac{\partial H^T}{\partial M} + \frac{dH_j^D}{dM})\frac{\partial M_{-j}}{\partial Q_C} \tag{4.18}$$

Conditions (18) inform us that the optimal carbon quota level for each nation is the one that equates the slope of the nation's iso-utility curve, $(p_{C_j} - \frac{\partial H^T}{\partial Q_C})/(\frac{\partial H^T}{\partial M} + \frac{dH_j^D}{dM})$ to the

[6]As stated in the introduction, SRM provision and reduction of carbon emissions are impure public goods in our model. They are imperfect strategic substitutes as equations (15) reveal. Our paper contributes to the public economics literature by considering the strategic interactions between these two impure public goods. It is also important to notice that the imperfect substitutability between the two impure public goods is not implied by our modeling assumptions with respect to the functional forms of damage and cost functions. It follows from the facts that are national-specific benefits associated with expansions in carbon emissions (consumer surplus produced by energy consumption) and national-specific costs associated with expansions in SRM provision (damages from droughts).

slope of the best response function of the other nation, $\frac{\partial M_{-j}}{\partial Q_C}$.

4.4.4 Game IV: Simultaneous Commitments on SRM Policies

Having observed $\{M_j, M_{-j}\}$, government j selects non-negative Q_{C_j} to maximize its indirect utility function (10) subject to (13a), taking Q_{C_j} as given, in the second stage. The first-order conditions are equations (15a). Differentiating each equation (15a) with respect to M yields

$$\frac{\partial Q_{C_j}}{\partial M} = \frac{-2h^T(\frac{\partial p_{C_{-j}}}{\partial Q_{C_{-j}}})}{\frac{\partial p_{C_1}}{\partial Q_{C_1}}\frac{\partial p_{C_2}}{\partial Q_{C_2}} - 2h^T(\frac{\partial p_{C_1}}{\partial Q_{C_1}} + \frac{\partial p_{C_2}}{\partial Q_{C_2}})} > 0 \tag{4.19}$$

where the sign of the each of the two equations in (19) follows from the facts that both the denominator and the numerator of the ratio in the right side of each equation (19) are positive - see Proposition 4.2. Note that each nation's carbon quota rises with the global SRM level at a rate that is less than one. In addition, it also follows that the global carbon emission level rises at a rate that is less than one when the global SRM level expands. Thus, an expansion in global SRM leads to a net decrease in the global temperature!

In the first stage, the first-order conditions are ($j = 1, 2$, $-j \neq j$):

$$-\frac{\partial H^T}{\partial M} = \frac{dK^M}{dM_j} + \frac{dH_j^D}{dM} + \frac{\partial H^T}{\partial Q_C}\frac{\partial Q_{C_{-j}}}{\partial M} \tag{4.20}$$

Equations (20) tell us that the optimal SRM level for each nation is the one at which the slope of the nation's indifference curve, $-(\frac{dK^M}{dM_j} + \frac{\partial H^T}{\partial M} + \frac{dH^D}{dM})$ is equal to the slope of the best response function of nation $-j$, $\frac{\partial H^T}{\partial Q_C}\frac{\partial Q_{C_{-j}}}{\partial M}$.

We can summarize the temperature outcomes of the sequential games as follows.

Proposition 4.3. *If both nations simultaneously commit to carbon permit policies, national SRM levels rise with the carbon quotas. The global temperature rises following each unit increase in the global carbon quota. If both nations simultaneously commit to SRM policies, national carbon quotas rise with national SRM levels. The global temperature falls* *following each unit increase in the global SRM level.*

4.4.5 Game V: National Leadership in Carbon Permit Policy

In this game, nation 1 is the leader in carbon policy. Nation 2 observes the carbon quota chosen by nation 1 and chooses its carbon quota in the second stage. In the third stage, both nations choose SRM levels simultaneously. In the fourth stage, both nations choose sulfur quotas simultaneously. Thus, this game adds one stage to the timing of game III. Relative to the policy setting captured by game III, it enables us to discover the effects associated with a nation being a leader in carbon policy.

Having observed Q_{C_1}, nation 2 chooses non-negative Q_{C_2} to maximize its payoff function subject to (13a) and (16). The first- and second-order conditions for the problem solved by government 2 are the same as in the first stage of game III. Hence, we can get $Q_{C_2}(Q_{C_1})$ from condition (19):

$$\frac{\partial Q_{C_2}}{\partial Q_{C_1}} = (\frac{\partial p_{C_2}}{\partial Q_{C_2}})/(\frac{\partial^2 v_2}{\partial Q_{C_2}^2}) - 1. \tag{4.21}$$

It is straightforward to show that $0 < (\frac{\partial p_{C_2}}{\partial Q_{C_2}})/(\frac{\partial^2 v_2}{\partial Q_{C_2}^2}) < 1$. Thus, equation (21) yields $-1 < \partial Q_{C_2}/\partial Q_{C_1} < 0$. Each unit increase in nation 1's carbon quota leads to a reduction in nation 2's carbon quota at a rate smaller than one. The net effect of each unit increase in nation 1's carbon quota is an increase in global carbon emissions and a subsequent increase in the global temperature.

Consider now the first stage. The first-order condition for an interior solution is:

$$p_{C_1} - [\frac{\partial H^T}{\partial Q_C} + (\frac{\partial H^T}{\partial M} + \frac{dH^D}{dM})](1 + \frac{\partial Q_{C_2}}{\partial Q_{C_1}}) = 0. \tag{4.22}$$

Condition (22) informs us that the optimal quota level for nation 1 is determined by the condition that equates $p_{C_1}/[\frac{\partial H^T}{\partial Q_C} + (\frac{\partial H^T}{\partial M} + \frac{dH^D}{dM})] - 1$, the slope of nation 1's indifference curve, to $\frac{\partial Q_{C_2}}{\partial Q_{C_1}}$, the slope of nation 2's best response function.

4.4.6 Game VI: National Leadership in SRM Policy

In this game, nation 1 is the leader in SRM policy. Nation 2 observes the SRM level provided by nation 1 and then chooses its SRM level in the second stage. In the third stage, carbon pollution quotas are chosen simultaneously. Finally, sulfur quotas are chosen simultaneously in the fourth stage.

In the second stage, nation 2 chooses $M_2 \geq 0$ to maximize its indirect utility, subject to (14a) and $Q_{C_j}(M)$. The first- and second-order conditions are the same as in the first stage of game IV. Then we can obtain $M_2(M_1)$ from condition (21). Hence,

$$\frac{\partial M_2}{\partial M_1} = -\frac{h^D(\delta_1 + \delta_2) + 2h^T(1 - \frac{\partial Q_C}{\partial M})(1 - \frac{\partial Q_{C_1}}{\partial M})}{2k^M + h^D(\delta_1 + \delta_2) + 2h^T(1 - \frac{\partial Q_C}{\partial M})(1 - \frac{\partial Q_{C_1}}{\partial M})} < 0. \tag{4.23}$$

Condition (23) reveals that SRM levels are strategic substitutes and that the rate of substitution is less than one in absolute value.

In the first stage, nation 1 determines M_1 to to maximize its indirect utility subject to the reactions in the subsequent stages. The first-order condition is

$$(\frac{\partial H^T}{\partial M} + \frac{dH_1^D}{dM} + \frac{\partial H^T}{\partial Q_C}\frac{\partial Q_{C_2}}{\partial M})(1 + \frac{\partial M_2}{\partial M_1}) + \frac{dK^M}{dM_1} = 0. \tag{4.24}$$

Condition (24) shows that the optimal SRM level for nation 1 is determined by the tangency condition which states that the slope of nation 1's indifference curve, $-\frac{dK^M}{dM_1}/(\frac{\partial H^T}{\partial M} + \frac{dH_1^D}{dM} + \frac{\partial H^T}{\partial Q_C}\frac{\partial Q_{C_2}}{\partial M}) - 1$, is equal to the slope of nation 2's best response function, $\frac{\partial M_2}{\partial M_1}$.

4.5 First-mover Advantage

We now demonstrate that a nation always benefits from being a policy leader, either in carbon policy or in SRM policy. Consider carbon policy first. By being a leader in carbon policy, nation 1 selects its optimal quantity in game V on firm 2's reaction curve. Thus, the choice nation 1 makes in game III (in which carbon policies are chosen simultaneously) is available to this nation when it makes its choice in game V. Since nation 1 selects a different carbon quantity quota in the equilibrium of game V and the equilibrium for this game is unique, nation 1 strictly prefers the carbon quota of game V to the carbon quota of game III. This is a revealed preference argument. Revealed preference and the fact that the equilibrium for game VI is unique also implies that nation 1 strictly prefers the choice it makes in game VI to the choice it makes in game IV. Thus, nation 1 strictly prefers to move first in each type of policy to moving simultaneously with nation 2.

We now demonstrate that nation 1 prefers to be the leader than to be the follower. As in Section 4.2.3, it is straightforward to show that

$$\partial Q_{C_1}/\partial Q_{C_2} < 0, \tag{4.25}$$

$$\partial Q_{M_1}/\partial Q_{M_2} < 0. \tag{4.26}$$

Result (25) informs us that carbon quotas are strategic substitutes. Result (26) informs us that SRM levels are strategic substitutes. Hence, we have the following facts:

$$\begin{cases} H^T &= h^T(C-M)^2 \\ C &= Q_{C_1} + Q_{C_2} \end{cases} \Rightarrow Q_{C_1} \text{ and } Q_{C_2} \text{ are technologically perfect substitutes } (i)$$

$$\begin{cases} H_j^D &= \delta_j h^D M^2 \\ M &= M_1 + M_2 \end{cases} \Rightarrow M_1 \text{ and } M_2 \text{ are technologically perfect substitutes } (ii)$$

Conditions (21) and (25) $\Rightarrow Q_{C_1}$ and Q_{C_2} are strategic substitutes. (iii)

Conditions (23) and (26) $\Rightarrow M_1$ and M_2 are strategic substitutes. (iv)

Proposition 4.4. *Given (i) and (iii), a nation always prefers to be the leader rather than the follower in carbon quota policy. Given (ii) and (iv), a nation always prefers to be the leader rather than the follower in SRM policy.*

Proof. It is a direct application of Varian's proof for the preference of quantity leadership in duopoly games. See [51], p. 297. $\square$

Together with the revealed preference results, which state that moving first is strictly preferable to moving simultaneously in each type of policy, we can affirm:

Proposition 4.5. *A nation always prefers to move first in either type of policy, carbon quota or SRM.*

4.6 Comparisons across Equilibria

Proposition 4.5 is important because it enables us to predict that, under our modeling assumptions, a nation will always attempt to be the leader in climate change policies. We are, however, unable to predict whether the leader has a policy preference. We now assign specific values to parameters of the utility, cost and harm functions in order to compare payoffs across equilibria and then to expand our predictions of play.

4.6.1 Symmetric Economy

We first assume that the economy is symmetric. In this case, drought and acid rain damage functions are identical. In the baseline case, let $a = 1/2, b = 1, k^C = k^S = k^M = k^E = h^A = h^T = h^D = 1/2, \delta_1 = \delta_2 = 1, \theta_1 = \theta_2 = 1$ and $\bar{x}_j = 0, j = 1, 2$. Solving consumer j's and producer j's maximization problems, the general equilibrium results in nation j are: $e_j = E_j = (1 + Q_{C_j} + Q_{S_j})/4$, $R_{C_j} = (1 - 3Q_{C_j} + Q_{S_j})/4$, $R_{S_j} = (1 + Q_{C_j} - 3Q_{S_j})/4$, $p_{e_j} = (3 - Q_{C_j} - Q_{S_j})/4$, $p_{C_j} = (1 - 3Q_{C_j} + Q_{S_j})/4$ and $p_{S_j} = (1 + Q_{C_j} - 3Q_{S_j})/4$, $j = 1, 2$. Plugging these functions into equation (10), we obtain the indirect utility function of nation j:

$$
\begin{aligned}
& v_j(Q_{C_j}, Q_{C_{-j}}, Q_{S_j}, M_j, M_{-j}) \\
=& [1 + 2(Q_{C_j} + Q_{S_j}) + 2Q_{C_j}Q_{S_j}]/8 - [4M_j^2 + 4M^2 + 4(Q_C - M)^2 + 3Q_{C_j}^2 + 7Q_{S_j}^2]/8.
\end{aligned}
\tag{4.27}
$$

It is useful to compute the first order condition of nation j's maximization problem with respect to Q_{S_j}, which holds in each game:

$$
\partial v_j / \partial Q_{S_j} = 0 \Rightarrow Q_{S_j} = (1 + Q_{C_j})/7, j = 1, 2. \tag{4.28}
$$

Solving the governments' maximization problems in the different games using the payoff functions (27) and conditions (28), we compute the Nash equilibrium outcomes.

The main results of the comparisons across equilibria in this baseline case are presented in Table 4.2.

Table 4.2 – Main results of baseline comparisons across equilibria.

Rankings	1	2	3	4	5	6
V	II	IV	VI	I	III	V
v_1	II	V	VI	IV	I	III
v_2	II	IV	I	VI	III	V
H_1^A	V	III	I	IV	VI	II
H_2^A	III	I	IV	VI	V	II
H^D	V	III	I	II	IV	VI
H^T	V	VI	IV	III	I	II

Proposition 4.6. *The rankings of payoffs and environmental damages in descending order (i.e., column 1 displays the highest values and column 6 displays the lowest values) are as follows:*[7]

Proof. It is available from the authors upon request. □

[7] In this table, we assume that nation 1 is the Stackelberg leader in games V and VI.

The results summarized in Proposition 4.6 are remarkable. Given the first-move advantage, the most likely non-cooperative scenarios are games V and VI, since these appear as the first and second most preferable options from the point of view of the leader, if we discard the cooperative game II as an option. If, in addition, we allow the leader to choose between leadership in carbon policy and leadership in SRM policy, the leader chooses carbon policy. Now, consider the ranking for global welfare. Not surprisingly, the highest level of global welfare is obtained at the equilibrium in which carbon and SRM policies are fully coordinated - game II. However, the other ranking positions provide us with very interesting messages for the inefficient allocations: (i) global welfare is higher when SRM policy is determined before carbon policy; (ii) for each type of policy commitment, global welfare is higher when there is no leadership; (iii) SRM policy commitments are superior to no policy commitment and the latter is superior to carbon policy commitments; (iv) SRM policy commitment without national leadership is second best; and (v) policy leadership in carbon policy is the worst scenario for the globe!

The first message follows from two comparisons: (a) between the global welfare levels implied by the equilibria for games IV and III; and (b) between global welfare levels implied by the equilibria for games VI and V. Games III and IV are three-stage games characterized by simultaneous choices in each stage. Since the equilibrium for game IV yields a higher level of global welfare than the equilibrium for game III, the globe should prefer a setting in which national authorities simultaneously choose SRM policies in the first stage to a setting in which national authorities simultaneously choose carbon policies in the first stage. Games V and VI are four-stage games characterized by individual choices in the first two stages and simultaneous choices in the last two stages. Since the equilibrium for game VI yields a higher level of global welfare than the equilibrium for game V, the globe should prefer a setting in which there is leadership in SRM policy to a setting in which there is policy leadership in carbon policy.

The second message follows from two comparisons: (a) between the global welfare levels implied by the equilibria for games III and V; and (b) between the global welfare levels implied by the equilibria for games IV and VI. The third message follows from the second message and two comparisons: (a) between the global welfare levels implied by the equilibria for games IV and I; and (b) between the global welfare levels implied by the equilibria for games I and III. The comparisons reveal that the global welfare level produced by the equilibrium game IV is higher than the global welfare level produced by the equilibrium for game I and the latter is higher than the global welfare level produced by the equilibrium for game III. The fourth message is straightforward, since the level of global welfare implied by the equilibrium for game IV is the highest among those produced by the inefficient equilibria. The last message is obvious.

For the inefficient allocations, one can understand the ranking for global welfare if in the comparisons we are able to rationalize the combined effects of policy leadership and ”easy riding” on the provision of SRM and on the setting of carbon quotas.[8] Not only the leader (nation 1) receives in the equilibrium for game V the highest payoff obtained by this nation in the set of inefficient games, the payoff received by nation 2 in the equilibrium for game V is the lowest obtained by this nation in the set of inefficient games. The equilibrium for game V features the highest global warming and drought damages, implying that the setting in which nation 1 is a policy leader in carbon policy yields the highest levels of carbon emission and SRM provision. To show this, note that Proposition 4.6 informs us that $H^{D^{V}} > H^{D^{III}} > H^{D^{I}} > H^{D^{II}} > H^{D^{IV}} > H^{D^{VI}}$. Since $H^D = M^2/2$, we have $M^V > M^{III} > M^I > M^{II} > M^{IV} > M^{VI}$. Proposition 4.6 also reveals that $H^{T^{V}} > H^{T^{VI}} > H^{T^{IV}} > H^{T^{III}} > H^{T^{I}} > H^{T^{II}}$. Since $H^T = (Q_C - M)^2/2$, we can combine the results to obtain $Q_C^V = \max\{Q_C^I, Q_C^{II}, Q_C^{III}, Q_C^{IV}, Q_C^V, Q_C^{VI}\}$.

The equilibrium for game V yields the highest levels of acid rain damage and sulfur dioxide emission in nation 1 - recall that $H_1^A = Q_{S_1}^2/2$. As clearly revealed by equation

[8]The term “easy riding” was first introduced in the literature by Cornes and Sandler (1984). This paper provides the first comprehensive analysis on easy rather than free riding in games that examine voluntary contributions to a public good.

(28), the carbon quota in nation 1 is an increasing function of the sulfur quota. Hence, the equilibrium for game V also features the highest level of carbon dioxide emission in nation 1. This setting, therefore, is characterized by overprovision of SRM (i.e., the amount of SRM provided is higher than the globally efficient amount obtained in the equilibrium for game II), the highest degree of easy riding on mitigation of carbon emissions and the highest levels of sulfur and carbon emissions in nation 1.

If we remove the policy leadership status of nation 1 in carbon policy, but still consider a setting in which carbon policy is determined before SRM policy, we are able to capture the effects promoted by leadership in carbon policy by comparing the national outcomes obtained in the equilibria for games III and V. The payoffs for nations 1 and 2 in the equilibrium for game III are the lowest and second lowest payoffs earned by these nations, respectively. Since the global warming and drought damages associated with the equilibrium of game III are smaller than their counterparts in the equilibrium for game V, removing policy leadership in carbon policy produces global benefits - both the degree of overprovision in SRM and the degree of easy riding in carbon mitigation are reduced. However, there are also national impacts associated with the removal of leadership in carbon policy - acid rain damages are reduced in nation 1 but increased in nation 2, implying that sulfur and carbon emissions in nation 1 are reduced, but sulfur and carbon emissions in nation 2 are increased.

Policy leadership in SRM, on the other hand, is more desirable from a global perspective. Not only the payoff that nation 1 receives in game VI is third best, the payoff earned by nation 2 in the equilibrium for game VI is the highest among the inefficient payoffs earned by this nation. The level of drought damage associated with the equilibrium for game VI is the lowest among all scenarios, implying that SRM is underprovided in the equilibrium for game VI. The level of acid rain damage faced by nation 1 in the equilibrium for game VI is the second lowest among all scenarios. This implies that the degrees of sulfur and carbon emission mitigation in nation 1 are second best. By comparing the equilibria for games

V and VI we can capture the effects of switching the type of policy leadership. Since the equilibrium for game V features overprovision of SRM and the highest degree of easy riding in carbon and sulfur mitigation in nation 1 and the equilibrium for game VI features suboptimal provision of SRM and second best degrees of mitigation of carbon and sulfur emissions, we see that the policy leader always chooses its policy instruments to minimize its contributions to improving the negative effects promoted by climate change.

The findings illustrated in Proposition 4.6 yield an interesting policy prescription other than forcing the nations to behave cooperatively. The policy prescription concerns the implementation of a global agreement on SRM policy in full anticipation that the nations will behave non-cooperatively. Provided this agreement leads the nations to make simultaneous commitments with respect to SRM policy, the resulting outcome will be second best for the globe. Even if the agreement is unable to prevent one nation from becoming a leader in SRM policy, the outcome will be superior to the most likely outcome in absence of the agreement ? game IV is third best for the globe.

4.6.2 Robustness: Asymmetric Drought and Acid Rain Damages

The numerical analysis in subsection 4.5.1 considers symmetric drought and acid rain damage functions. It also assumes identical values for most parameters. In this section, we check whether the rankings of payoffs are robust to differences in drought damages, acid rain damages and changes in the values of some parameters. Table 4.3 below shows the results of some mathematical simulations.

The most important results of our analysis in the symmetric case concerned our predictions with respect to the “choice of a scenario” that a policy leader will make and the contrasting (non-cooperative) choice that the globe will make if it can exercise this option. The leader prefers the scenario that arises under game V. The non-cooperative choice of the globe is scenario IV. If the globe is unable to prevent a nation from becoming a leader in policymaking, then it prefers scenario VI.

Table 4.3 reveals that these conclusions remain true even in the presence of asymmetric drought and acid rain damages. The table reports results for four situations in which the drought damage in nation 1 is larger than the drought damage in nation 2: (i) when the relative damage is ten times larger in nation 1; (ii) when the relative damage is four times larger in nation 1; (iii) when the relative damage is two times larger in nation 1; and (iv) when the relative damage is 1.33 times larger in nation 1 than in nation 2. In each situation, we consider the effects of incremental changes in theta 1 and theta 2. To understand this, first row for the global payoff V. We use the baseline values for the parameters other than delta 1, delta 2, theta 1 and theta 2. We then start by considering how the ranking displayed in Table 4.2 changes (if there is any change at all) when we fix theta 1 equal to 0.1 and let theta 2 to be equal to 0.2. We compute the results and move to the next iteration, where theta 1 is still kept equal to 0.1 and theta 2 is increased to 0.3, and so on until theta 2 is equal to 1. After this, we set theta 1 equal to 0.2 and let theta 2 to be equal to 0.1. The next iteration keeps the value of theta 1 constant and increases the value of theta 2 to 0.2, and so on until theta 2 equals 1. The computations end when both theta 1 and theta 2 equal 1. The results in the first row demonstrate that the baseline ranking for the global payoff remains unchanged in the four situations for all possible combinations of theta 1 and theta 2 values.

The second row of the table shows that the baseline ranking for the global payoff remains unchanged under the first three asymmetric-drought-damage situations even when one of the taste parameters, a, is evaluated in the range [0.05, 0.7] and the other parameters are kept constant at their baseline values. The baseline ranking for the global payoff changes in the last asymmetric-drought-damage situation. For 0.05 incremental changes in a, the ranking changes because the relative positions of scenarios I and VI alternate (i.e., for some a values in the interval [0.05, 0.7], V is larger in scenario I than in scenario VI, but for some other a values the opposite is true. Alternating pairwise rankings are also observed for different values of b (another taste parameter) in the interval [1, 10] in the first two asymmetric-drought-damage situations. For k^S (a technological parameter) in [0.05, 1], the baseline ranking for the global payoff remains unchanged. We could have offered similar conclusions for the other types of parameters regarding the broad range of parameter values under which the baseline ranking for the global payoff remains unchanged, but decided to keep the table small and display the results for “representative” taste and technological parameters.

When we consider the potential changes in the baseline rankings for the payoffs earned by the leader and the follower, we notice that there are not as many circumstances under which the baseline ranking for each type of player remains unchanged as in the baseline ranking for the global payoff. However, the message that the leader prefers scenario V among all non-cooperative scenarios remains unchanged. In addition, restricting our attention to the most likely non-cooperative scenarios V and VI, the follower still prefers scenario VI to V.

4.7 Conclusion

This paper represents an initial exercise on the effects promoted by policy commitments on competing instruments designed to reduce the negative effects associated with global warming. We consider a global economy consisting of two nations and in which production of energy generates both sulfur and carbon emissions and the nations can use carbon quotas and SRM provision to reduce the negative effects associated with global warming. Although solar radiation management is a global pure public good with respect to climate change, its provision has two types of monetary costs (namely, technological and drought damage), which formally make it an impure public good. We show that a nation always views carbon and sulfur quotas as strategic complements. We also show that a nation always prefers to be a policy leader, irrespective if the leadership is in carbon policy or SRM policy. For various values of parameters of utility and technology, we can demonstrate that, among the inefficient scenarios, a nation prefers to be a policy leader in carbon policy, but the globe prefers a setting in which SRM policy is simultaneously determined by the competing nations before these nations simultaneously determine carbon policy. If the globe is faced with situations in which a nation displays policy leadership and it can choose between policy leadership in carbon and SRM policies, it will choose policy leadership in SRM. From the globe's perspective, if a fully coordinated agreement is unavailable, it prefers the settings in which SRM policy is determined before carbon policy.

The low cost alternative produced by SRM provision relative to mitigation of carbon emissions leads us to believe that nations will engage in the provision of SRM in the near future. Due to questions of national security and sovereignty, nations may not be forthcoming in disclosing key information about their activities related to development of SRM. In future work, we plan to incorporate uncertainty and asymmetric information in national provision of SRM into the model and study the predictions of play of imperfectly informed governments. Another interesting avenue for future work is to utilize the recent developments in aggregative games to examine extensions of our model to a general setting with a large number of nations.[9]

[9] See, e.g., [18] and [15].

Table 4.3 – Rankings with asymmetric drought and acid rain damages.

Payoffs	Parameter Values	$(\delta_1, \delta_2) = (1, 0.1)$	$(\delta_1, \delta_2) = (1, 0.25)$	$(\delta_1, \delta_2) = (1, 0.5)$	$(\delta_1, \delta_2) = (1, 0.75)$
V	Baseline	Robust	Robust	Robust	Robust
	$a \in [0.05, 0.7]$	Robust	Robust	Robust	Alternate I & VI
	$b \in [1, 10]$	Alternate IV & VI	Alternate IV & VI	Robust	Robust
	$k^S \in [0.05, 1]$	Robust	Robust	Robust	Robust
v_1	Baseline	Alternate II & V	Robust	Robust	Robust
	$a \in [0.1, 1]$	Alternate II & V	Alternate II & V	Alternate II & V	Robust
	$b \in [1, 10]$	Alternate (II & V) or (III & IV)	Alternate (II & V) or (III & IV)	Alternate (II & V) or (III & IV)	Alternate (II & V) or (III & IV)
	$k^S \in [0.05, 1]$	Alternate II & V	Robust	Robust	Robust
v_2	Baseline	Change VI & I	Change VI & I	Alternate VI & I	Robust
	$a \in [0.15, 1]$	Change VI & I	Change VI & I	Alternate VI & I	Alternate VI & I
	$b \in [1, 10]$	Change VI & I	Change VI & I	Alternate VI & I	Robust
	$k^S \in [0.05, 1]$	Change VI & I	Change VI & I	Alternate VI & I	Robust

Appendices

APPENDIX A

APPENDIX FOR CHAPTER 2

A.1 Proofs

A.1.1 Proof of Proposition 2.1

Proof. We first observe that for any player $j \neq i$ such that $x_j = 0$, $g_{ij} = 0$ since $k > 0$. This means $\sum_{j \in N_i^{out}} x_j = |N_i^{out}|$. Then $\Pi_i^{ow}(x_i = 1) = f(1 + |N_i^{out}|) - c - k|N_i^{out}| + b_i(|N_i^{in}|)$, $\Pi_i^{ow}(x_i = 0) = f(|N_i^{out}|) - k|N_i^{out}|$. Recall that $\check{y} = \arg\max_y [f(y) - ky]$.

If $\sum_{j \neq i} x_j \geq \check{y}$, then $\Pi_i^{ow}(x_i = 0) = f(\check{y}) - k\check{y}$, where $|N_i^{out}| = \check{y}$; $\Pi_i^{ow}(x_i = 1) = f(\check{y}) - c - k(\check{y} - 1) + b_i(|N_i^{in}|) = [f(\check{y}) - k\check{y}] + [b_i(|N_i^{in}|) - (c - k)]$, where $|N_i^{out}| = \check{y} - 1$. Therefore, $\Delta\Pi_i^{ow} \equiv \Pi_i^{ow}(x_i = 1) - \Pi_i^{ow}(x_i = 0) = b_i(|N_i^{in}|) - (c - k) > 0$ if and only if $|N_i^{in}| \geq \lceil b_i^{-1}(c - k) \rceil$.

If $\sum_{j \neq i} x_j \leq \check{y} - 1$, then $\Pi_i^{ow}(x_i = 0) = f(\sum_{j \neq i} x_j) - k \sum_{j \neq i} x_j$, where $|N_i^{out}| = \sum_{j \neq i} x_j$; $\Pi_i^{ow}(x_i = 1) = f(1 + \sum_{j \neq i} x_j) - c - k \sum_{j \neq i} x_j + b_i(|N_i^{in}|)$, where $|N_i^{out}| = \sum_{j \neq i} x_j$. Therefore, $\Delta\Pi_i^{ow} \equiv \Pi_i^{ow}(x_i = 1) - \Pi_i^{ow}(x_i = 0) = f(1 + \sum_{j \neq i} x_j) - f(\sum_{j \neq i} x_j) + b_i(|N_i^{in}|) - c = f'(\sum_{j \neq i} x_j) - c + b_i(|N_i^{in}|)$. Recall that $\hat{y} = \arg\max_y [f(y) - cy] = \max\{y \in \mathbb{N} : f'(y - 1) \geq c\} \leq \check{y} - 1$. When $\sum_{j \neq i} x_j \leq \hat{y} - 1$, $\Delta\Pi_i^{ow} > 0$ always; when $\hat{y} \leq \sum_{j \neq i} x_j \leq \check{y} - 1$, $\Delta\Pi_i^{ow} > 0$ if and only if $|N_i^{in}| \geq \lceil b_i^{-1}[c - f'(\sum_{j \neq i} x_j)] \rceil$.

Hence, in an equilibrium $s^* = (x^*, g^*)$, either i) $\hat{y} \leq \sum_{i \in N} x_i^* \leq \check{y}$, and every volunteer receives an incoming link from everyone else; or ii) $\sum_{i \in N} x_i^* \geq \check{y} + 1$, and every player gets access to $\check{y}$ units of public good by forming links to the volunteers. Further note that if $\sum_{i \in N} x_i^* \geq \check{y}$, then players' best responses are not strict. This implies that in every strict NE, $\hat{y} \leq \sum_{i \in N} x_i^* \leq \check{y}$. $\square$

A.1.2 Proof of Proposition 2.2

Proof. We begin with deriving a player's best-response function. Given others' strategies, if player i chooses to free ride, her payoff will be $\Pi_i^{tw}(x_i = 0) = f(\sum_{j\in N_i} x_j) - k|N_i^{out}| \leq f(\sum_{j\in N_i^{in}} x_j + \sum_{j\in N_i^{out}} x_j) - k|N_i^{out}|$, where the equality holds if $N_i^{out} \cap N_i^{in} = \emptyset$. Meanwhile, since $k > 0$, then a free rider will only link to volunteers, and so $\sum_{j\in N_i^{out}} = |N_i^{out}|$. Then $\Pi_i^{tw}(x_i = 0) = f(\sum_{j\in N_i} x_j) - k\sum_{j\in N_i^{out}} x_j = f(\sum_{j\in N_i} x_j) - k\sum_{j\in N_i} x_j + k\sum_{j\in N_i^{in}} x_j$. Recall that $\check{y} = \arg\max_y[f(y) - ky]$. Therefore, if $\sum_{j\neq i} x_j \geq \check{y}$, $\Pi_i^{tw}(x_i = 0) = f(\check{y}) - k\check{y} + k\sum_{j\in N_i^{in}} x_j$, where $|N_i^{out}| = \check{y} - \sum_{j\in N_i^{in}} x_j$; if $\sum_{j\neq i} x_j \leq \check{y} - 1$, $\Pi_i^{tw}(x_i = 0) = f(\sum_{j\neq i} x_j) - k\sum_{j\neq i} x_j + k\sum_{j\in N_i^{in}} x_j$, where $|N_i^{out}| = \sum_{j\neq i} x_j - \sum_{j\in N_i^{in}} x_j$.
If player i chooses to volunteer, then her payoff will be $\Pi_i^{tw}(x_i = 1) = f(1 + \sum_{j\in N_i} x_j) - c - k|N_i^{out}| + b_i(|N_i|) = f(1 + \sum_{j\in N_i^{in}} x_j + \sum_{j\in N_i^{out}} x_j) - c - k|N_i^{out}| + b_i(|N_i^{in}| + |N_i^{out}|)$. Let $N_{i,V}^{in}$ ($N_{i,F}^{in}$) denote the set of links from volunteers (free riders) to player i and let $N_{i,V}^{out}$ ($N_{i,F}^{out}$) denote the set of links from player i to volunteers (free riders). Then $\Pi_i^{tw}(x_i = 1) = f(1 + |N_{i,V}^{in}| + |N_{i,V}^{out}|) - c - k(|N_{i,V}^{out}| + |N_{i,F}^{out}|) + b_i(|N_{i,V}^{in}| + |N_{i,F}^{in}| + |N_{i,V}^{out}| + |N_{i,F}^{out}|)$. Note that for volunteer i, an outgoing link to another volunteer will bring at least as many benefits as to a free rider, and so volunteer i will form links with free riders only after she has formed links with all other volunteers. This leads us to first determine $|N_{i,V}^{out}|$ and then $|N_{i,F}^{out}|$. Recall that $\hat{y} = \arg\max_y[f(y) - cy]$, $\check{y} = \arg\max_y[f(y) - ky]$ and $\tilde{m}_i = \arg\max_m[b_i(m) - km]$.

(1) If $\sum_{j\neq i} x_j \leq \check{y} - 1$, then volunteer i forms links to every other unlinked volunteer, *i.e.*, $|N_{i,V}^{out}| = \sum_{j\neq i} x_j - |N_{i,V}^{in}|$. Then $|N_i^{in}| + |N_{i,V}^{out}| = |N_{i,V}^{in}| + |N_{i,F}^{in}| + |N_{i,V}^{out}| = |N_{i,F}^{in}| + |N_{i,V}| = |N_{i,F}^{in}| + \sum_{j\neq i} x_j$.

i) If $|N_i^{in}| + |N_{i,V}^{out}| \leq \min\{\tilde{m}_i, n-1\}$, then $|N_{i,F}^{out}| = \min\{\tilde{m}_i, n-1\} - |N_i^{in}| - |N_{i,V}^{out}| = \min\{\tilde{m}_i, n-1\} - |N_{i,F}^{in}| - \sum_{j\neq i} x_j$ and $|N_i| = \min\{\tilde{m}_i, n-1\}$. Then $\Pi_i^{tw}(x_i = 1) = f(1 + \sum_{j\neq i} x_j) - c - k(\min\{\tilde{m}_i, n-1\} - |N_i^{in}|) + b_i(\min\{\tilde{m}_i, n-1\})$. This gives $\Delta\Pi_i^{tw} \equiv \Pi_i^{tw}(x_i = 1) - \Pi_i^{tw}(x_i = 0) =$

$[f'(\sum_{j\neq i} x_j) - c] + [b_i(\min\{\tilde{m}_i, n-1\}) - k\min\{\tilde{m}_i, n-1\}] + k(|N^{in}_{i,F}| + \sum_{j\neq i} x_j) \leq [f'(\sum_{j\neq i} x_j) - c] + b_i(\min\{\tilde{m}_i, n-1\}) - k$. If $\sum_{j\neq i} x_j \leq \hat{y} - 1$, then $\Delta\Pi^{tw}_i > 0$ always.

ii) Otherwise, $|N^{out}_{i,F}| = 0$. Then $\Pi^{tw}_i(x_i = 1) = f(1+\sum_{j\neq i} x_j) - c - k(\sum_{j\neq i} x_j - |N^{in}_{i,V}|) + b_i(|N^{in}_i| + \sum_{j\neq i} x_j - |N^{in}_{i,V}|) = f(1+\sum_{j\neq i} x_j) - c - k(\sum_{j\neq i} x_j - |N^{in}_{i,V}|) + b_i(|N^{in}_{i,F}| + \sum_{j\neq i} x_j)$. This gives $\Delta\Pi^{tw}_i \equiv \Pi^{tw}_i(x_i = 1) - \Pi^{tw}_i(x_i = 0) = [f'(\sum_{j\neq i} x_j) - c] + b_i(|N^{in}_{i,F}| + \sum_{j\neq i} x_j) \geq [f'(\sum_{j\neq i} x_j) - c] + b_i(\tilde{m}_i)$. Again, if $\sum_{j\neq i} x_j \leq \hat{y} - 1$, then $\Delta\Pi^{tw}_i > 0$ always.

(2) If $\sum_{j\neq i} x_j \geq \check{y}$, then form outgoing links until the sum of the marginal social and consumption benefits an additional link brings becomes less than the marginal cost k.

The above derivations show that in every NE, $\sum_{i\in N} x^*_i \geq \hat{y}$.

(1) When $\hat{y} \leq \sum_{i\in N} x^*_i \leq \check{y}$, every volunteer's provision is pivotal to others, and so every volunteer is minimally linked with everyone else, and a volunteer may sponsor a link to a free rider when her marginal social benefit is high ($i.e.$, $\tilde{m}_i$ is large). This leads to a core-periphery structure in which the hub players are volunteers. Note that the best response of every player is strict. This gives type-I strict NE.

(2) When $\sum_{i\in N} x^*_i \geq \check{y}+1$, a volunteer's provision may or may not be pivotal to another player. Depending on the extent of her marginal social benefit, a volunteer may or may not be minimally linked with everyone else and a free rider will sponsor at most $\check{y}$ number of links. In a strict NE, since a free rider is indifferent among which volunteers to link with, then a link between a volunteer and a free rider must go from the former to the latter. Meanwhile, recall that a volunteer prefers forming a link with another volunteer to linking with a free rider and $\tilde{m}_i = \arg\max_m[b_i(m) - km]$. Then to ensure that there exist volunteers who still want to sponsor links with free riders after being minimally linked with all volunteers, it must be that $\sum_{i\in N} x^*_i \leq$

$\max_{i\in N}\tilde{m}_i$. In addition, since a volunteer is indifferent among which free riders to link with, then for any volunteer i with $|N_{i,F}^{out}| \geq 1$, we have $|N_{i,V}| = \sum_{j\neq i} x_j - 1$, $|N_{i,F}^{out}| = n - \sum_{j\neq i} x_j$ and $|N_i| = n-1$. When players are heterogeneous in their social benefits from each recipient, this could lead to three groups of volunteers, say groups A, B and C. Every group A volunteer links with every free rider and $\bar{m}_A > \tilde{m}_A \geq n-1$; every group B volunteer is minimally linked with all group A volunteers and forms links with all group C volunteers, and $\tilde{m}_B \leq n-2, \bar{m}_B \geq \sum_{i\in N} x_i^* - 1$; every group C volunteer receives links from volunteers in groups A and B, but does not form any outgoing links, and $\tilde{m}_C < \bar{m}_C \leq \sum_{i\in N} x_i^* - 2$. This implies that $\max_{i\in N}\tilde{m}_i = n-1$. We further note that the existence of group A volunteers is essential for satisfying free riders' need for the public good and $\sum_{i\in A} x_i^* \geq \check{y}$. The existence of the other two groups depends on the distribution of the social benefits from each recipient. This gives type-II strict NE.

□

A.1.3 Proof of Corollary 2.1

Proof. We note that the conclusion is immediate for strict NE of one-way flow of benefits and type-I strict NE of two-way flow of benefits, since $\hat{y} \leq \sum_{i\in N} x_i^* \leq \check{y}$ where $\check{y}$ is a fixed constant.

For type-II strict NE of two-way flow of benefits, recall from the proof of Proposition 2 that the existence of such equilibria requires that there are group-A volunteers with $\tilde{m}_i \geq n-1, \forall i \in A$ for every free rider to have access to the public good. Since $\tilde{m}_i$ is some fixed constant for any player i, then the condition is unlikely to hold when $n \to \infty$. This implies that when $n \to \infty$, every limit strict NE belongs to type I and so the law of the few holds.

□

A.1.4 Proof of Lemma 2.1

Proof. In the absence of social incentives, $b_i(\cdot) = 0, \tilde{m}_i = 0, \forall i$. For either type of flow, every player forms outgoing links just for access to the public good provided by others. Then based on the proofs of Proposition 1 and 2, for any player i, when $\sum_{j \neq i} x_j \leq \hat{y} - 1$, $\Delta\Pi_i^{ow} = \Delta\Pi_i^{tw} = f'(\sum_{j \neq i} x_j) - c$; when $\sum_{j \neq i} x_j \geq \hat{y}$, $\Delta\Pi_i^{ow} = \Delta\Pi_i^{tw} = k - c < 0$. Since $\hat{y} = \max\{y \in \mathbb{N} : f'(y-1) \geq c\}$, then $\sum_{i \in N} x_i^* = \hat{y}$ in every NE. $\square$

A.1.5 Proof of Proposition 2.3

Proof. We first prove the results regarding the number of equilibrium volunteers. When $\check{y} = \hat{y}$, the result follows directly from Proposition 1. When $\check{y} \geq \hat{y} + 1$, to study whether the presence of social incentives could lead to more volunteers in every strict NE, consider the strategy profile in which there exist $\hat{y}$ number of volunteers. We want to know if an incumbent free rider has an incentive to deviate and starts volunteering. Note that the current payoff of free rider i is $\Pi_i^{ow}(x_i = 0) = f(\hat{y}) - k\hat{y}$. Given others' strategies, if she deviates unilaterally and starts volunteering, then her payoff becomes $\Pi_i^{ow}(x_i^{dev} = 1) = f(\hat{y} + 1) - c - k\hat{y}$. So the potential gain from deviation $\Delta\Pi_i^{ow}(x_i = 0, x_i^{dev} = 1) = f'(\hat{y}) - c < 0$. This shows that no matter how large the social incentives are, a strict NE with $\hat{y}$ number of volunteers always exists.

Second, we observe from above that an incumbent free rider's potential gain from deviation does not depend on her social benefits from each recipient since she does not receive any incoming link. Therefore, a player may choose to free ride just because of zero incoming link, rather than low social benefits from each recipient. Then any player could be a volunteer in a strict NE. $\square$

A.1.6 Proof of Proposition 2.4

Proof.

Part 1. We first prove the results regarding the number of equilibrium volunteers. As in the proof of Proposition 3, we again consider the strategy profile in which there exist $\hat{y}$ number of volunteers and study the deviation incentive of an incumbent free rider. For free rider i, her current payoff is $\Pi_i^{tw}(x_i = 0) = f(\hat{y}) - k(\hat{y} - |N_i^{in}|)$. Given others' strategies, if she deviates unilaterally and starts volunteering, then she will still be minimally linked with all other volunteers as $\check{y} \geq \hat{y} + 1$, and she may form links with other free riders. Then her payoff becomes $\Pi_i^{tw}(x_i^{dev} = 1) = f(1+\hat{y}) - c - k(\hat{y} - |N_i^{in}| + |N_{i,F}^{out}|) + b_i(\hat{y} + |N_{i,F}^{out}|)$. Here, $|N_{i,F}^{out}| = \min\{\tilde{m}_i, n-1\} - \hat{y}$ if $\hat{y} \leq \min\{\tilde{m}_i, n-1\}$; $|N_{i,F}^{out}| = 0$ otherwise. Therefore, the potential gain from deviation is

$$\Delta\Pi_i^{tw}(x_i = 0, x_i^{dev} = 1) = \begin{cases} f'(\hat{y}) - c + b_i(\hat{y}) \text{ if } \hat{y} > \min\{\tilde{m}_i, n-1\}, \\ f'(\hat{y}) - c + b_i(\tilde{m}_i) - k(\tilde{m}_i - \hat{y}) \text{ if } \hat{y} \leq \tilde{m}_i \leq n-1, \\ f'(\hat{y}) - c + b_i(n-1) - k(n-1-\hat{y}) \text{ if } \hat{y} \leq n-1 \leq \tilde{m}_i. \end{cases}$$

One can see that given other parameter values, $\Delta\Pi_i^{tw}(x_i = 0, x_i^{dev} = 1) > 0$ could hold in every subcase if the marginal social incentive $b_i'(\cdot)$ is sufficiently large. Therefore, if the marginal social incentives are sufficiently large for at least $\hat{y} + 1$ number of players, then it is certain that there exist at least $\hat{y} + 1$ number of volunteers in every strict NE. This proves the "if" part of the statement.

Conversely, to ensure that there exist at least $\hat{y} + 1$ number of volunteers in every strict NE, there must exist player i whose marginal social incentives are sufficiently large so that $\Delta\Pi_i^{tw}(x_i = 0, x_i^{dev} = 1) > 0$ always. In addition, there must be at least $\hat{y} + 1$ number of them so that even when all $\hat{y}$ number of the incumbent volunteers are those with sufficiently high social incentives, there will always be another player with sufficiently high social incentives to join the volunteer set. This proves the "only if" part of the statement.

Part 2. We now prove the results for comparing the social benefits from each recipient of a volunteer and those of a free rider. As in Proposition 3, whenever a volunteer has more incoming links than a free rider does, then we could not conclude on which one enjoys higher social benefits from each recipient. This happens when we examine a volunteer-free rider pair in a type-I strict NE or when the volunteer belongs to group C in a type-II strict NE with both group B and C volunteers present. We show below that in the rest of the cases, every volunteer enjoys higher social benefits from each recipient than a free rider does, which all happens when the strict NE is of type II and the number of incoming links a free rider has is more than that of a volunteer.

Consider a type-II strict NE of two-way flows with $|V|$ number of volunteers. Let player A be a representative volunteer of group A, player B be a representative volunteer of group B, player C be a representative volunteer of group C, and player D be a representative free rider. Let $|V_A|$ be the size of group A, $|V_B|$ be the size of group B and $|V_C|$ be the size group C. By the definition of strict NE, the following holds.

$$
\begin{aligned}
V_A :& f(|V|) - c - k(n-1-|N_A^{in}|) + b_A(n-1) \\
& > \begin{cases} f(\check{y}) - k(\check{y} - |N_A^{in}|) \text{ if } |N_A^{in}| \leq \check{y}, \\ f(|N_A^{in}|) \text{ if } |N_A^{in}| \geq \check{y}+1; \end{cases} \\
V_B :& f(|V|) - c - k(|V|-1-|N_B^{in}|) + b_B(|V|-1) \\
& > \begin{cases} f(\check{y}) - k(\check{y} - |N_B^{in}|) \text{ if } |N_B^{in}| \leq \check{y}, \\ f(|N_B^{in}|) \text{ if } |N_B^{in}| \geq \check{y}+1; \end{cases} \\
V_C :& f(|V_A|+|V_B|+1) - c + b_C(|V_A|+|V_B|) > f(|V_A|+|V_B|); \\
F_D :& f(|V_A|) > \max_{|N_{D,V}^{out}|,|N_{D,F}^{out}|} [f(|V_A|+1+|N_{D,V}^{out}|) - c - k(|N_{D,V}^{out}|+|N_{D,F}^{out}|) \\
& \quad + b_D(|V_A|+|N_{D,V}^{out}|+|N_{D,F}^{out}|)].
\end{aligned}
$$

We first compare $b'_B(m)$ with $b'_D(m)$, for given $m \in \mathbb{N}$. Since the inequality for F_D above holds for any $|N_{D,V}^{out}|$ and $|N_{D,F}^{out}|$, we set $|N_{D,V}^{out}| = |V| - |V_A| - 1$ and $|N_{D,F}^{out}| = 0$.

Then we have the following for F_D: $f(|V_A|)-k|V_A| > f(|V|)-c-k(|V|-1)+b_D(|V|-1)$. The inequalities for V_B give: if $|N_B^{in}| \leq \check{y}$, then $f(|V|) - c - k(|V| - 1) + b_B(|V| - 1) > f(\check{y})-k\check{y}$; if $|N_B^{in}| \geq \check{y}+1$, then $f(|V|)-c-k(|V|-1)+b_B(|V|-1) > f(|N_B^{in}|)-k|N_B^{in}|$. Also recall that $|V_A| \geq \check{y}$, $|N_{B,V}^{in}| \leq |V_A|$ and $f(y) - ky$ is increasing in y when $y \leq \check{y}$ and decreasing in y when $y \geq \check{y} + 1$. Therefore, if $|N_B^{in}| \leq \check{y}$, then $f(|V|) - c - k(|V| - 1) + b_B(|V|-1) > f(\check{y})-k\check{y} > f(|V_A|)-k|V_A| > f(|V|)-c-k(|V|-1)+b_D(|V|-1)$, and so $b'_B(m) > b'_D(m), \forall m \in \mathbb{N}$. If $|N_B^{in}| \geq \check{y}+1$, then $f(|V|)-c-k(|V|-1)+b_B(|V|-1) > f(|N_B^{in}|) - k|N_B^{in}| \geq f(|V_A|) - k|V_A| > f(|V|) - c - k(|V| - 1) + b_D(|V| - 1)$, and so $b'_B(m) > b'_D(m), \forall m \in \mathbb{N}$.

Now recall that in a type-II strict NE, based on the proof of Proposition 2, $\tilde{m}_A > \tilde{m}_B$ and so $b'_A(m) > b'_B(m) > b'_D(m), \forall m \in \mathbb{N}$.

Finally, we compare $b'_C(m)$ with $b'_D(m)$, for given $m \in \mathbb{N}$. We note that if $\bar{m}_D \leq |V_A|$, then $|N_{D,V}^{out}| = |N_{D,F}^{out}| = 0$. Then the inequality for F_D gives: $c - b_D(|V_A|) > f'(|V_A|)$. Meanwhile, the inequality for V_C gives: $f'(|V_A| + |V_B|) > c - b_C(|V_A| + |V_B|)$. Since $f'(|V_A|) \geq f'(|V_A| + |V_B|)$, then $b_D(|V_A|) < b_C(|V_A| + |V_B|)$. If $|V_B| = 0$, then we have $b'_C(m) > b'_D(m), \forall m \in \mathbb{N}$.

□

A.1.7 Proof of Remark 2.1

We provide the proof of the following claim that is used in the arguments of the remark. The rest of the arguments are easy to see.

Claim A.1. *If the cost of every transition into a state is lower than that of every transition out of it, then that state is the unique SSE.*

Proof. Suppose that the cost of every transition into state s is lower than that of every transition out of it. Let s' be another state. We note that the two states have the same number of trees rooted at them. Meanwhile, in a tree $T_{s'}$ rooted at state s', s reaches s' either directly or indirectly.

First, consider the case in which s reaches s' directly. There exists an edge from s to s' and a path from every state $s'' \neq s, s'$ to s'. Now consider the trees rooted at state s. There must be one tree T_s with an edge from s' to s and all other paths from s'' to s' are the same as those in the tree $T_{s'}$. Since the cost of the transition from s' to s is lower than that of the transition from s to s', then the exponential cost of the tree T_s[1] is lower than that of the tree $T_{s'}$.

Second, consider the case in which s reaches s' indirectly. There exists an edge from s to $s'' \neq s, s'$ and a path from s'' to s'. Again, consider the trees rooted at state s. There must be one tree T_s with an edge from s' to s and all other paths from s'' to s' are the same as those in the tree $T_{s'}$. Since the cost of the transition from s' to s is lower than that of the transition from s to s'', then the exponential cost of the tree T_s is lower than that of the tree $T_{s'}$.

Therefore, for any tree rooted at a state $s' \neq s$, there always exists a tree rooted at s with a lower exponential cost. This implies that state s is the unique SSE. $\square$

A.1.8 Proof of Proposition 2.5

Proof. When the flow of benefits is one-way, Proposition 1 says that every strict NE has a core-periphery structure. Let $NE_{(1)}$ be a strict NE of one-way flows and $\hat{y} \leq |V_{(1)}| \leq \check{y}-1$. It is easy to see that there always exists another strict NE, $NE_{(2)}$, such that $V_{(1)} \subsetneq V_{(2)}$ and $\hat{y}+1 \leq |V_{(2)}| \leq \check{y}$. We will show below that $\Pi_{i,(1)} < \Pi_{i,(2)}, \forall i \in N$. Before proving this result, we note that $f(|V_{(1)}|) - k|V_{(1)}| < f(|V_{(2)}|) - k|V_{(2)}|$ since $\hat{y} \leq |V_{(1)}| < |V_{(2)}| \leq \check{y}$.

- $\forall i \in V_{(1)}, \Pi_{i,(1)}(x_i = 1) = f(|V_{(1)}|) - c - k(|V_{(1)}| - 1) + b_i(n-1) < f(|V_{(2)}|) - c - k(|V_{(2)}| - 1) + b_i(n-1) = \Pi_{i,(2)}(x_i = 1)$;
- $\forall i \in V_{(2)} \backslash V_{(1)}, \Pi_{i,(1)}(x_i = 0) = f(|V_{(1)}|) - k|V_{(1)}| < f(|V_{(1)}|) - c - k(|V_{(1)}| - 1) + b_i(n-1) = \Pi_{i,(2)}(x_i = 1)$;

[1] Appendix B provides details about the exponential cost of a tree.

- $\forall i \in F_{(2)}, \Pi_{i,(1)}(x_i = 0) = f(|V_{(1)}|) - k|V_{(1)}| < f(|V_{(2)}|) - k|V_{(2)}| = \Pi_{i,(2)}(x_i = 0)$.

Therefore, $\Pi_{i,(1)} < \Pi_{i,(2)}, \forall i \in N$. This implies that $NE_{(2)}$ Pareto dominates $NE_{(1)}$.

□

A.1.9 Proof of Proposition 2.6

Proof.

Part 1. Let $NE_{[1]}$ be a type-I strict NE and $\hat{y} \leq |V_{[1]}| \leq \check{y} - 1$. We will first show that there always exists another type-I strict NE, $NE_{[2]}$, such that $V_{[1]} \subsetneq V_{[2]}$, with all links in $NE_{[2]}$ preserving the same directions as in $NE_{[1]}$ and all additional links in $NE_{[2]}$ going from free riders to volunteers. Then we will show that such $NE_{[2]}$ Pareto dominates $NE_{[1]}$.

Note that for any player $i \notin V_{[1]}$, in the presence of $|V| \in [|V_{[1]}|, \check{y}]$ number of volunteers and $|N^{in}_{i,V}|$ number of incoming links from existing volunteers, i's payoff when volunteering is $\Pi_i(x_i = 1) = f(|V| + 1) - c - k(|V| - |N^{in}_{i,V}|) + b_i(n-1)$ and her payoff when free riding is $\Pi_i(x_i = 0) = f(|V|) - k(|V| - |N^{in}_{i,V}|)$. Then $\Pi_i(x_i = 1) > \Pi_i(x_i = 0)$ and i will choose to volunteer always. Moreover, since the linking structures are preserved for all volunteers in $NE_{[1]}$, and each of them could access more units of public goods now, then they will also continue volunteering. This shows that $NE_{[2]}$ always exists.

Next, we show that $NE_{[2]}$ as defined above Pareto dominates $NE_{[1]}$.

- $\forall i \in V_{[1]}$, $\Pi_{i,[1]}(x_i = 1) = f(|V_{[1]}|) - c - k(|V_{[1]}| - 1 - |N^{in}_{i,V}|) + b_i(n-1) < f(|V_{[2]}|) - c - k(|V_{[2]}| - 1 - |N^{in}_{i,V}|) + b_i(n-1) = \Pi_{i,[2]}(x_i = 1)$;

- $\forall i \in V_{[2]} \backslash V_{[1]}, \Pi_{i,[1]}(x_i = 0) = f(|V_{[1]}|) - k(|V_{[1]}| - |N^{in}_{i,V}|) < f(|V_{[2]}|) - c - k(|V_{[2]}| - 1 - |N^{in}_{i,V}|) + b_i(n-1) = \Pi_{i,[2]}(x_i = 1)$;

- $\forall i \in F_{[2]}$, $\Pi_{i,[1]}(x_i = 0) = f(|V_{[1]}|) - k(|V_{[1]}| - |N^{in}_{i,V}|) < f(|V_{[2]}|) - k(|V_{[2]}| - |N^{in}_{i,V}|) = \Pi_{i,[2]}(x_i = 0)$.

Therefore, $\forall i \in N$, $\Pi_{i,[2]} > \Pi_{i,[1]}$, which implies that $NE_{[2]}$ Pareto dominates $NE_{[1]}$.

Part 2. Now we compare the efficiency between a type-I strict NE and a type-II strict NE. We first illustrate with an example that players who volunteer in both type-I and type-II strict NE could be better off in the former equilibrium with less volunteers. Consider $n = 9$, $f(y) = \sqrt{y}$, $c = \frac{1}{2\sqrt{2}}$, $k = \frac{1}{2\sqrt{3}}$, $b_i(m) = b_i\sqrt{m}$ and $b_1 = b_2 = b_3 = \frac{5\sqrt{2}}{\sqrt{3}}$, $b_4 = ... = b_9 = \frac{1}{\sqrt{3}}$. Then $\hat{y} = 2$, $\check{y} = 3$, $\tilde{m}_1 = \tilde{m}_2 = \tilde{m}_3 = 50$, $\tilde{m}_4 = ... = \tilde{m}_9 = 1$. One can check that there exist type-I strict NE with two or three volunteers, and a unique type-II strict NE in which players with $b_i = 50$ are group-A volunteers and those with $b_i = 1$ are group-C volunteers. Now we compare the payoffs of player i who volunteer in both a type-I strict NE (denoted by $NE_{[5]}$) and the type-II strict NE (denoted by $NE_{[6]}$). We consider $NE_{[5]}$ in which the three players with $b_i = 50$ are volunteers and the chosen player i receives incoming links from everyone else. In $NE_{[6]}$, player i forms outgoing links to all six free riders and receives links from the other two volunteers (this yields the highest payoff for player i in a type-II strict NE). Then $\Pi_{i,[5]} - \Pi_{i,[6]} = [f(3) - c + b_i(n-1)] - [f(8) - c - 6k + b_i(n-1)] = 2(\sqrt{3} - \sqrt{2}) > 0$. Hence, player i is strictly better off in the type-I strict NE. We note that such a player i who volunteers in both types of strict NE always exists. This is because any group-A volunteer in a type-II NE enjoys high social benefits from each recipient of the public goods and will sponsor connections with any unlinked players no matter how many others volunteer. Therefore, the example implies that type-I strict NE do not Pareto dominate type-II strict NE in general.

Now we check the reverse direction of Pareto dominance. It is easy to see that for players who choose to free ride in both NE, they will always be better off with accessing more public goods. Therefore, if a type-I strict NE Pareto dominates a type-II strict NE, then it must be that all free riders in the former are volunteers in the latter, which could not hold in general. Therefore, type-I strict NE also do not Pareto dominate type-II strict NE in general. $\square$

A.1.10 Proof of Proposition 2.7

Proof. When the flow of benefits is one-way, we know that every player forms outgoing links so as to access the public good provided by others. Based on the proof of Proposition 1, the maximum number of links player i will sponsor is $\tilde{y}_i$. Therefore, when everyone has a strict best response, it must be that $\sum_{i \in N} x_i^* \leq \min_{i \in N} \tilde{y}_i$. In addition, we know that for any player i, if $\sum_{j \neq i} x_j \leq \hat{y}_i - 1$, then her best response is always to volunteer. Therefore, in equilibrium, $\sum_{i \in N} x_i^* \geq \min_{i \in N} \hat{y}_i$. Together we have $\min_{i \in N} \hat{y}_i \leq \sum_{i \in N} x_i^* \leq \min_{i \in N} \tilde{y}_i$ in every strict NE. □

A.1.11 Proof of Proposition 2.8

Proof. When the flow of benefits is two-way, while a free rider forms outgoing links only for access to the public good provided by others, a volunteer links with others for both consumption benefits of the public good and social benefits of volunteering. Based on the proof of Proposition 2, we know that for any player i, if $\sum_{j \neq i} x_j \leq \hat{y}_i - 1$, then her best response is to volunteer always. This means in every NE, $\sum_{i \in N} x_i^* \geq \min_{i \in N} \hat{y}_i$.

Now we note that for free rider i, if $\sum_{j \neq i} x_j \geq \tilde{y}_i$, then F_i will form links with at most $\tilde{y}_i$ number of randomly picked volunteers. Therefore, in a strict NE, when $\sum_{i \in N} x_i^* \leq \min_{i \in F} \tilde{y}_i$, every volunteer will be minimally linked with all free riders; when $\sum_{i \in N} x_i^* \geq \min_{i \in F} \tilde{y}_i$, then there exists free rider j such that not every volunteer's provision is pivotal to her consumption benefits of the public good and she will form at most $\tilde{y}_j$ number of links to the volunteers.

Regarding volunteer i, from the proof of Proposition 2, we know that in a strict NE, if $|N_{i,F}^{out}| \geq 1$, then $|N_i| = n - 1$ (since a volunteer prefers linking with another volunteer to linking with a free rider, and a volunteer is indifferent among linking with all other volunteers, or all free riders, respectively); if $|N_{i,V}^{out}| \geq 1$, then $|N_{i,V}| = |V| - 1$ (since a volunteer is indifferent among linking with all other volunteers). This means that in a strict NE, some volunteers are minimally linked with all other $n - 1$ players, some are minimally linked with all other volunteers but do not form links with free riders, and the rest only receive links from others.

Putting these together, we obtain the following for the shape of strict NE. i) When $\sum_{i \in N} x_i^* \leq \min_{i \in N} \tilde{y}_i$, every volunteer's provision is pivotal to everyone else and so g^* has a core-periphery structure in which volunteers lie at the core and links between free riders and volunteers can go in both directions. ii) When $\sum_{i \in N} x_i^* \geq \min_{i \in F} \tilde{y}_i$, g^* is a multipartite graph in which links between volunteers and free riders always go from the former to the latter, and free riders (and some of the volunteers) form an independent set. iii) When $\min_{i \in V} \tilde{y}_i \leq \sum_{i \in N} x_i^* \leq \min_{i \in F} \tilde{y}_i$, there might exist some volunteers such that not every other volunteer's provision is pivotal to their consumption benefits of the public good and their valuations on the social benefits are not high. Then those players form an independent set in which they only receive links from all others outside the set. Therefore, g^* is a multipartite graph in which every volunteer is minimally linked with all free riders while volunteers may or may not be minimally linked with each other.

□

A.1.12 Proofs of Propositions 2.9 and 2.10

Proof. In the proofs of Proposition 1 and 2, for player i's decision making, attach δ to all the terms associated with $x_j, \forall j \neq i$. Then the conclusions follow easily. □

A.2 Dynamic Setup for Section 2.5.1

Time is discrete and indexed by $t = 1, 2, \ldots$. The strategy profile $s^t \in S$ represents the state of the play at time t. The dynamic process is as follows. At each time period t, the system randomly selects a player i to revise her strategy based on the current play of others, which generates a Markov chain on S. To avoid the system being locked into strict NE, we introduce noises[2] and assume that player i updates her strategy following the logit quantal

[2]For example, players may make mistakes or experiment with new strategies occasionally.

response rule in myatt082 and myatt081. Let $\epsilon > 0$ be a noise parameter. Then

$$\log \frac{\mathbb{P}\{s_i^{t+1} = s_i^{M_h} | s^t = s\}}{\mathbb{P}\{s_i^{t+1} = s_i^{BR} | s^t = s\}} = -\frac{\Delta\Pi_i^{M_h}(s)}{\epsilon},$$

where s_i^{BR} represents player i's best response to s_{-i}, $s_i^{M_h}$ is the h^{th} possible mutated strategy (in the presence of noises, there are many possible mutations since a player makes one volunteering and $n-1$ linking decisions simultaneously) and $\Delta\Pi_i^{M_h}(s) \equiv \Pi_i(s_i^{t+1} = s_i^{BR}|s^t = s) - \Pi_i(s_i^{t+1} = s_i^{M_h}|s^t = s)$ is the corresponding loss in payoff. Then

$$\mathbb{P}\{s_i^{t+1} = s_i^{M_h} | s^t = s\} = \frac{e^{-\frac{\Delta\Pi_i^{M_h}(s)}{\epsilon}}}{1 + \sum_{h'} e^{-\frac{\Delta\Pi_i^{M_{h'}}(s)}{\epsilon}}}.$$

The logit responses approximate best replies when $\epsilon \to 0$. The Markov chain is now irreducible and aperiodic, and has a unique stationary distribution μ^ϵ showing how often each state is played in the long run. Let $\mu \equiv \lim_{\epsilon\to 0} \mu^\epsilon$ be the limit distribution. A state is *stochastically stable* if its limit probability in μ is bounded above zero.

The long-run play depends on the rate at which transition probabilities vanish as $\epsilon \to 0$. As in myatt081, we define such a rate as the *exponential cost* $\mathcal{E}$ of a probability. More formally, let $p(\epsilon)$ be a continuous function on $\mathbb{R}^+ \cup \{\infty\}$. Then either $\mathcal{E} = \infty$ if $p(\epsilon) = 0, \forall\epsilon > 0$, or the limit $\mathcal{E} = -\lim_{\epsilon\to 0} \epsilon \log p(\epsilon)$ exists. This means that $p(\epsilon)$ behaves as $e^{-\mathcal{E}/\epsilon}$ does when $\epsilon \to 0$. In our setup, since the dynamics follows the logit quantal responses, one can verify that the exponential cost of a mutation equals its loss in payoff, $i.e.$, $\mathcal{E}(s_i^{M_h}|s) \equiv \mathcal{E}(\mathbb{P}\{s_i^{t+1} = s_i^{M_h}|s^t = s\}) = \Delta\Pi_i^{M_h}(s)$. Moreover, the exponential cost of the product of mutation probabilities equals the sum of individual exponential costs, $i.e.$, $\mathcal{E}(p_1(\epsilon)p_2(\epsilon)) = \mathcal{E}(p_1(\epsilon)) + \mathcal{E}(p_2(\epsilon))$.

We apply the method of rooted trees to find the stochastically stable states. Think of each state s as a node of a directed graph on S. Let $|S|$ be the number of states (or nodes) in the set S. Then a *rooted tree at* s is a spanning tree T_s on S such that $\forall s' \neq s$, there is a unique directed path from s' to s. Note that for the system to transit from state s' to s, there are many possible routes of costly mutations. In a tree T_s, define the exponential cost of a directed edge (s', s), $\mathcal{E}_{s's}$, as the lowest cost of these routes. Then the exponential cost of the tree T_s equals the sum of the exponential costs of its edges, $i.e.$, $\mathcal{E}_s = \sum_{(s',s)} \mathcal{E}_{s's}$. myatt081 shows that the states with minimum-exponential-cost rooted trees are stochastically stable. Meanwhile, as $\epsilon \to 0$, to characterize the limiting stationary distribution, we only need to look at smaller trees where each node corresponds to a different strict pure NE.

APPENDIX B

APPENDIX FOR CHAPTER 3

B.1 Proofs

B.1.1 Proof of Result 3.3

Proof. From Result 1, we know that in the absence of FTAs, every player would be better off in structure III (IV) compared to I. However, structure I is stable unless all three countries deviate and add links between each other. Therefore, whenever I, III and IV have interior solutions, all of them are NE and BS, but only III and IV are GS and CPNE. □

B.1.2 Proof of Result 3.4

Proof. Let's first look at the situation in which structures I and IV have interior solutions. From Result 2, we know that every country would be better off in IV. However, structure I is stable unless all three countries deviate and add links between each other. Therefore, both I and IV are NE and BS while only IV is GS and CPNE.

Now let's consider the situation in which structures II and IV have interior solutions. According to Result 2, we know that club members in II would prefer II to IV while the reverse holds for the non-member country. Therefore, the club members in II would never want to add links to the non-member country and there would not be profitable deviation from II. On the other hand, two of the three countries in IV are likely to deviate together by deleting links to the other country. Hence, both II and IV are NE, but only II is BS, GS and CPNE.

We continue to look at the situation in which structures III and IV have interior solutions. From Result 2, we know that the hub country in III prefers III to IV, while the reverse holds for the spoke countries. This implies that no subset of countries would profit from deleting links and deviating from IV, and so IV is stable by all criteria. As for the deviation from III, we know that both spoke countries would be better off by adding a link between them although deviating alone would not be profitable. Therefore, both III and IV are NE, but only IV is BS, GS and CPNE.

Finally, we examine the situation in which structures I, II and IV have interior solutions. Recall that in Result 2, club members in II would prefer II to I and IV. Therefore, two of the three countries in I would deviate together by forming a link with each other; and two of the three countries in IV would deviate together by deleting links to the other country. Hence, although I, II and IV are all NE, only II is BS, GS and CPNE. □

B.1.3 Proof of Result 3.8

Proof. When structures V, VII and VIII have interior solutions, from Result 6, we know that $w_2^V = w_3^V < w_2^{VIII} = w_3^{VIII}$, $w_3^{VII} < w_2^{VIII}$ and $w_2^{VII} < w_1^{VIII}$. Therefore, in structures V and VII, countries 2 and 3 would have an incentive to form a club between each other. This implies that structures V and VII are not BS. Meanwhile, we observe that when c is not very low (in our simulation, this means $c \geq 0.3$), $w_4^{VIII} < w_3^{VII} < w_3^{VIII} < w_1^{VII}$, and so countries 3 and 4 in structure VIII would deviate together by withdrawing from the clubs with country 1 and form a club between themselves instead, which implies that structure VIII is not BS as well.

When structures IV, IX, X and XI have interior solutions, from Result 6, we know that $w_1^X < w_1^{IV} < w_1^{IX} = w_1^{XI} < w_3^X$. Therefore, country 3 in structure XI would have an incentive to withdraw from the club with country 4; country 1 in structure X would want to withdraw from the club with country 2; and countries 1 and 3, 1 and 4, 2 and 3, 2 and 4 in structure IV would deviate together by forming two clubs. This implies that structures IV, IX are NE, BS and CPNE; and structure IX is GS.

When structures IV, X and XI have interior solutions, from the welfare ranking $w_1^X < w_1^{IV} < w_1^{XI} < w_3^X$, we see that country 3 in XI would deviate by withdrawing from the club with country 4; countries 1 and 2 in structure X would deviate together by withdrawing from the clubs between countries 1 and 2, 1 and 3, and 2 and 4; and countries in IV would deviate together by forming clubs between 1 and 3, 1 and 4, 2 and 3, and 2 and 4. Therefore, IV is NE, BS and CPNE; X is NE; and XI is not NE.

When structures I, II, X and XI have interior solutions, from Result 6, we know that $w_1^X < w_1^I < w_3^{II} < w_1^{II} < w_1^{XI} < w_3^X$. Therefore, country 3 in structure XI would deviate by withdrawing from the club with 4; countries 1 and 2 in structure X would deviate together by withdrawing from the clubs with 3 and 4; countries 1, 2 and 3 in structure II would deviate together by forming clubs between 1 and 3, 1 and 4, 2 and 3, 2 and 4, 3 and 4; and countries 1 and 2 in structure I would deviate together by forming a club with each other. This means that structures I, II and X are NE, structure II is BS and CPNE, and none of them is GS.

When structures I, X and XI have interior solutions, from Result 6, we know that $w_1^X < w_1^I < w_1^{XI} < w_3^X$. Then country 3 in structure XI would deviate by withdrawing from the club with 4, countries 1 and 2 in structure X would deviate together by withdrawing from all the clubs they are involved in; and countries in structure I would deviate by forming clubs among each other. This implies that structures I and X are NE, structure I is BS and CPNE, and none of them is GS.

When structures V and VII have interior solutions, from Result 6, we know that $w_2^V < w_3^{VII} < w_1^{VII} < w_1^V$. Then countries 2 and 3 in structure V would deviate together by withdrawing from the club between 1 and 2, and forming a club between themselves. This implies that structures V and VII are NE, and structure VII is BS, CPNE and GS.

When structures V, VII, VIII and XI have interior solutions, from Result 6, we know that $w_2^V < w_4^{VIII} < w_3^{VII} < w_2^{VIII} < w_1^{XI} < w_1^{VII} < w_1^{VIII} < w_1^V$. Then in structures V and VII, countries 2 and 3 would have an incentive to form a club between each other; countries 3 and 4 in structure VIII would deviate together by withdrawing from the clubs with country 1 and form a club between themselves instead. Therefore, structures V, VII, VIII and XI are NE; and XI is BS, GS and CPNE.

When structures V, VII, VIII, IX and XI have interior solutions, from Result 6, we know that $w_2^V < w_4^{VIII} < w_3^{VII} < w_2^{VIII} < w_1^{IX} = w_1^{XI} < w_1^{VII} < w_1^{VIII} < w_1^V$. Based on the above analysis, structures V, VII, VIII, IX and XI are NE; IX and XI are BS, GS and CPNE.

When structures V, VII, IX and XI have interior solutions, from Result 6, we know that $w_2^V < w_3^{VII} < w_1^{IX} = w_1^{XI} < w_1^{VII} < w_1^V$. Then country 4 in structure V would want to deviate by withdrawing from the club with country 1 and forming another club with country 2; and countries 3 and 4 in VII would deviate together by forming a club between them. Therefore, structure V, VII, IX and XI are NE; structures IX and XI are BS are GS and CPNE.

When structures V, VII and XI have interior solutions, from Result 6, we know that $w_2^V < w_3^{VII} < w_1^{XI} < w_1^{VII} < w_1^V$. Then country 4 in structure V would want to deviate by withdrawing from the club with country 1 and forming another club with country 2; countries in VII would deviate together by forming clubs between 1 and 4, 2 and 3, and 3 and 4. Therefore, structures V, VII and XI are NE; structures VII and XI are BS; and structure XI is GS and CPNE. $\square$

B.2 Tables

Table B.1 – Rankings of national welfare in the absence of FTAs ($n = 3$).

c	θ	Welfare Ranking
0.1	0.01 to 2.52	no interior solutions
	2.53 to 2.76	$w_1^{III} = w_2^{III} = w_1^{IV}$
	2.77 to 4.08	$w_1^{I} < w_1^{III} = w_2^{III} = w_1^{IV}$
	4.09 to 5.11	w_1^{I}
	5.12 to 10	no interior solutions
0.3	0.01 to 2.72	no interior solutions
	2.73 to 3.43	$w_1^{III} = w_2^{III} = w_1^{IV}$
	3.44 to 4.08	$w_1^{I} < w_1^{III} = w_2^{III} = w_1^{IV}$
	4.09 to 5.17	w_1^{I}
	5.18 to 10	no interior solutions
0.5	0.01 to 2.8	no interior solutions
	2.81 to 3.73	$w_1^{III} = w_2^{III} = w_1^{IV}$
	3.74 to 4.08	$w_1^{I} < w_1^{III} = w_2^{III} = w_1^{IV}$
	4.09 to 5.21	w_1^{I}
	5.22 to 10	no interior solutions
0.7	0.01 to 2.84	no interior solutions
	2.85 to 3.91	$w_1^{III} = w_2^{III} = w_1^{IV}$
	3.92 to 4.08	$w_1^{I} < w_1^{III} = w_2^{III} = w_1^{IV}$
	4.09 to 5.24	w_1^{I}
	5.25 to 10	no interior solutions
0.9	0.01 to 2.87	no interior solutions
	2.88 to 4.03	$w_1^{III} = w_2^{III} = w_1^{IV}$
	4.04 to 4.08	$w_1^{I} < w_1^{III} = w_2^{III} = w_1^{IV}$
	4.09 to 5.26	w_1^{I}
	5.27 to 10	no interior solutions

Table B.2 – Rankings of national welfare in the presence of FTAs ($n = 3$).

c	θ	Welfare Ranking
0.1	0.01 to 1.33	no interior solutions
	1.34 to 1.82	$w_2^{III} < w_1^{III}$
	1.83 to 1.92	no interior solutions
	1.93 to 2.43	w_1^{IV}
	2.44 to 2.76	$w_3^{II} < w_1^{IV} < w_1^{II}$
	2.77 to 3.1	$w_3^{II} < w_1^{I} < w_1^{IV} < w_1^{II}$
	3.11 to 5.11	$w_1^{I} < w_1^{IV}$
	5.12 to 10	w_1^{IV}
0.3	0.01 to 1.77	no interior solutions
	1.78 to 1.92	$w_2^{III} < w_1^{III}$
	1.93 to 2.07	$w_2^{III} < w_1^{IV} < w_1^{III}$
	2.08 to 3.23	w_1^{IV}
	3.24 to 3.43	$w_3^{II} < w_1^{IV} < w_1^{II}$
	3.44 to 3.45	$w_3^{II} < w_1^{I} < w_1^{IV} < w_1^{II}$
	3.46 to 5.17	$w_1^{I} < w_1^{IV}$
	5.18 to 10	w_1^{IV}
0.5	0.01 to 1.91	no interior solutions
	1.92	$w_2^{III} < w_1^{III}$
	1.93 to 2.17	$w_2^{III} < w_1^{IV} < w_1^{III}$
	2.18 to 3.73	w_1^{IV}
	3.74 to 5.21	$w_1^{I} < w_1^{IV}$
	5.22 to 10	w_1^{IV}
0.7	0.01 to 1.92	no interior solutions
	1.93 to 1.98	w_1^{IV}
	1.99 to 2.22	$w_2^{III} < w_1^{IV} < w_1^{III}$
	2.23 to 3.91	w_1^{IV}
	3.92 to 5.24	$w_1^{I} < w_1^{IV}$
	5.25 to 10	w_1^{IV}
0.9	0.01 to 1.92	no interior solutions
	1.93 to 2.02	w_1^{IV}
	2.03 to 2.26	$w_2^{III} < w_1^{IV} < w_1^{III}$
	2.27 to 4.03	w_1^{IV}
	4.04 to 5.26	$w_1^{I} < w_1^{IV}$
	5.27 to 10	w_1^{IV}

Table B.3 – Equilibria in the absence of FTAs ($n = 3$).

c	θ	NE	BS	GS	CPNE
0.1	0.01 to 2.52	n/a	n/a	n/a	n/a
	2.53 to 2.76	III, IV	III, IV	III, IV	III, IV
	2.77 to 4.08	I, III, IV	I, III, IV	III, IV	III, IV
	4.09 to 5.11	I	I	I	I
	5.12 to 10	n/a	n/a	n/a	n/a
0.3	0.01 to 2.72	n/a	n/a	n/a	n/a
	2.73 to 3.43	III, IV	III, IV	III, IV	III, IV
	3.44 to 4.08	I, III, IV	I, III, IV	III, IV	III, IV
	4.09 to 5.17	I	I	I	I
	5.18 to 10	n/a	n/a	n/a	n/a
0.5	0.01 to 2.8	n/a	n/a	n/a	n/a
	2.81 to 3.73	III, IV	III, IV	III, IV	III, IV
	3.74 to 4.08	I, III, IV	I, III, IV	III, IV	III, IV
	4.09 to 5.21	I	I	I	I
	5.22 to 10	n/a	n/a	n/a	n/a
0.7	0.01 to 2.84	n/a	n/a	n/a	n/a
	2.85 to 3.91	III, IV	III, IV	III, IV	III, IV
	3.92 to 4.08	I, III, IV	I, III, IV	III, IV	III, IV
	4.09 to 5.24	I	I	I	I
	5.25 to 10	n/a	n/a	n/a	n/a
0.9	0.01 to 2.87	n/a	n/a	n/a	n/a
	2.88 to 4.03	III, IV	III, IV	III, IV	III, IV
	4.04 to 4.08	I, III, IV	I, III, IV	III, IV	III, IV
	4.09 to 5.26	I	I	I	I
	5.27 to 10	n/a	n/a	n/a	n/a

Table B.4 – Equilibria in the presence of FTAs ($n = 3$).

c	θ	NE	BS	GS	CPNE
0.1	0.01 to 1.33	n/a	n/a	n/a	n/a
	1.34 to 1.82	III	III	III	III
	1.83 to 1.92	n/a	n/a	n/a	n/a
	1.93 to 2.43	IV	IV	IV	IV
	2.44 to 2.76	II,IV	II	II	II
	2.77 to 3.1	I,II,IV	II	II	II
	3.11 to 5.11	I,IV	I,IV	IV	IV
	5.12 to 10	IV	IV	IV	IV
0.3	0.01 to 1.77	n/a	n/a	n/a	n/a
	1.78 to 1.92	III	III	III	III
	1.93 to 2.07	III,IV	IV	IV	IV
	2.08 to 3.23	IV	IV	IV	IV
	3.24 to 3.43	II,IV	II	II	II
	3.44 to 3.45	I,II,IV	II	II	II
	3.46 to 5.17	I,IV	I,IV	IV	IV
	5.18 to 10	IV	IV	IV	IV
0.5	0.01 to 1.91	n/a	n/a	n/a	n/a
	1.92	III	III	III	III
	1.93 to 2.17	III,IV	IV	IV	IV
	2.18 to 3.73	IV	IV	IV	IV
	3.74 to 5.21	I,IV	I,IV	IV	IV
	5.22 to 10	IV	IV	IV	IV
0.7	0.01 to 1.92	n/a	n/a	n/a	n/a
	1.93 to 1.98	III	III	III	III
	1.99 to 2.22	III,IV	IV	IV	IV
	2.23 to 3.91	IV	IV	IV	IV
	3.92 to 5.24	I,IV	I,IV	IV	IV
	5.25 to 10	IV	IV	IV	IV
0.9	0.01 to 1.92	n/a	n/a	n/a	n/a
	1.93 to 2.02	III	III	III	III
	2.03 to 2.26	III,IV	IV	IV	IV
	2.27 to 4.03	IV	IV	IV	IV
	4.04 to 5.26	I,IV	I,IV	IV	IV
	5.27 to 10	IV	IV	IV	IV

Table B.5 – Rankings of national welfare in the absence of FTAs ($n = 4$).

c	θ	Welfare Ranking
0.1	0.01 to 1.72	no interior solutions
	1.73 to 2.01	$w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	2.02 to 2.35	$w_1^{IV} < w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	2.36 to 3.16	$w_1^I < w_1^{IV} < w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	3.17 to 4.09	$w_1^I < w_1^{IV}$
	4.1 to 4.51	w_1^I
	4.52 to 10	no interior solutions
0.3	0.01 to 1.87	no interior solutions
	1.88 to 2.37	$w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	2.38 to 2.91	$w_1^{IV} < w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	2.92 to 3.16	$w_1^I < w_1^{IV} = w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	3.17 to 4.1	$w_1^I < w_1^{IV}$
	4.11 to 4.55	w_1^I
	4.56 to 10	no interior solutions
0.5	0.01 to 1.93	no interior solutions
	1.94 to 2.54	$w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	2.55 to 3.16	$w_1^{IV} < w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	3.17 to 4.16	$w_1^I < w_1^{IV}$
	4.17 to 4.57	w_1^I
	4.58 to 10	no interior solutions

Table B.5 (continued).

c	θ	Welfare Ranking
0.7	0.01 to 1.96	no interior solutions
	1.97 to 2.64	$w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	2.65 to 3.15	$w_1^{IV} < w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	3.16 to 3.3	w_1^{IV}
	3.31 to 4.16	$w_1^I < w_1^{IV}$
	4.17 to 4.59	w_1^I
	4.6 to 10	no interior solutions
0.9	0.01 to 1.99	no interior solutions
	2 to 2.71	$w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	2.72 to 3.16	$w_1^{IV} < w_1^V = w_1^{VII} = w_1^{VIII} = w_1^{IX} = w_1^X = w_1^{XI}$
	3.17 to 3.3	w_1^{IV}
	3.31 to 4.16	$w_1^I < w_1^{IV}$
	4.17 to 4.59	w_1^I
	4.6 to 10	no interior solutions

Table B.6 – Rankings of national welfare in the presence of FTAs ($n = 4$).

c	θ	Welfare Ranking
0.1	0.01 to 0.15	$w_2^V < w_3^{VII} < w_4^{VIII} < w_2^{VIII} < w_1^{VII} < w_1^{VIII} < w_1^V$
	0.16 to 0.53	$w_3^{VII} < w_4^{VIII} < w_2^{VIII} < w_1^{VII} < w_1^{VIII}$
	0.54 to 0.7	$w_4^{VIII} < w_2^{VIII} < w_1^{VIII}$
	0.71 to 1.24	no interior solutions
	1.25 to 1.33	w_1^{XI}
	1.34 to 1.41	$w_1^{IX} = w_1^{XI}$
	1.42	$w_1^X < w_1^{IX} = w_1^{XI} < w_3^X$
	1.43 to 2.04	$w_1^X < w_1^{IV} < w_1^{IX} = w_1^{XI} < w_3^X$
	2.05 to 2.09	$w_1^X < w_1^{IV} < w_1^{XI} < w_3^X$
	2.1 to 2.35	$w_1^X < w_1^{XI} < w_3^X$
	2.36 to 2.38	$w_1^X < w_1^I < w_1^{XI} < w_3^X$
	2.39 to 2.44	$w_1^X < w_1^I < w_3^{II} < w_1^{II} < w_1^{XI} < w_3^X$
	2.45 to 2.77	$w_1^X < w_1^I < w_1^{XI} < w_3^X$
	2.78 to 4.51	$w_1^X < w_1^{XI}$
	4.52 to 10	w_1^{XI}
0.3	0.01 to 0.65	no interior solutions
	0.66 to 0.77	$w_2^V < w_1^V$
	0.78 to 0.8	$w_2^V < w_3^{VII} < w_1^{VII} < w_1^V$
	0.81 to 1.05	$w_2^V < w_4^{VIII} < w_3^{VII} < w_2^{VIII} < w_1^{VII} < w_1^{VIII} < w_1^V$
	1.06 to 1.16	$w_4^{VIII} < w_3^{VII} < w_2^{VIII} < w_1^{VII} < w_1^{VIII}$
	1.17	$w_4^{VIII} < w_2^{VIII} < w_1^{VIII}$
	1.18 to 1.25	no interior solutions
	1.26 to 1.35	w_1^{XI}
	1.36 to 1.65	$w_1^{IX} = w_1^{XI}$
	1.66 to 1.73	$w_1^X < w_1^{IX} = w_1^{XI} < w_3^X$
	1.74 to 2.04	$w_1^X < w_1^{IV} < w_1^{IX} = w_1^{XI} < w_3^X$
	2.05 to 2.16	$w_1^X < w_1^{IV} < w_1^{XI} < w_3^X$
	2.17 to 2.77	$w_1^X < w_1^{XI} < w_3^X$
	2.78 to 2.91	w_1^{XI}
	2.92 to 4.55	$w_1^I < w_1^{XI}$
	4.56 to 10	w_1^{XI}

Table B.6 (continued).

c	θ	Welfare Ranking
0.3	0.01 to 0.65	no interior solutions
	0.66 to 0.77	$w_2^V < w_1^V$
	0.78 to 0.8	$w_2^V < w_3^{VII} < w_1^{VII} < w_1^V$
	0.81 to 1.05	$w_2^V < w_4^{VIII} < w_3^{VII} < w_2^{VIII} < w_1^{VII} < w_1^{VIII} < w_1^V$
	1.06 to 1.16	$w_4^{VIII} < w_3^{VII} < w_2^{VIII} < w_1^{VII} < w_1^{VIII}$
	1.17	$w_4^{VIII} < w_2^{VIII} < w_1^{VIII}$
	1.18 to 1.25	no interior solutions
	1.26 to 1.35	w_1^{XI}
	1.36 to 1.65	$w_1^{IX} = w_1^{XI}$
	1.66 to 1.73	$w_1^X < w_1^{IX} = w_1^{XI} < w_3^X$
	1.74 to 2.04	$w_1^X < w_1^{IV} < w_1^{IX} = w_1^{XI} < w_3^X$
	2.05 to 2.16	$w_1^X < w_1^{IV} < w_1^{XI} < w_3^X$
	2.17 to 2.77	$w_1^X < w_1^{XI} < w_3^X$
	2.78 to 2.91	w_1^{XI}
	2.92 to 4.55	$w_1^I < w_1^{XI}$
	4.56 to 10	w_1^{XI}
0.5	0.01 to 0.96	no interior solutions
	0.97 to 1.02	$w_2^V < w_1^V$
	1.03 to 1.06	$w_2^V < w_4^{VIII} < w_2^{VIII} < w_1^{VII} < w_1^{VIII} < w_1^V$
	1.07 to 1.24	$w_2^V < w_4^{VIII} < w_3^{VII} < w_2^{VIII} < w_1^{VII} < w_1^{VIII} < w_1^V$
	1.25 to 1.27	$w_2^V < w_4^{VIII} < w_3^{VII} < w_2^{VIII} < w_1^{XI} < w_1^{VII} < w_1^{VIII} < w_1^V$
	1.28 to 1.3	$w_2^V < w_3^{VII} < w_1^{XI} < w_1^{VII} < w_1^V$
	1.31 to 1.36	$w_3^{VII} < w_1^{XI} < w_1^{VII}$
	1.37	$w_3^{VII} < w_1^{IX} = w_1^{XI} < w_1^{VII}$
	1.38 to 1.75	$w_1^{IX} = w_1^{XI}$
	1.76 to 1.84	$w_1^X < w_1^{IX} = w_1^{XI} < w_3^X$
	1.85 to 2.04	$w_1^X < w_1^{IV} < w_1^{IX} = w_1^{XI} < w_3^X$
	2.05 to 2.19	$w_1^X < w_1^{IV} < w_1^{XI} < w_3^X$
	2.2 to 2.77	$w_1^X < w_1^{XI} < w_3^X$
	2.78 to 3.16	w_1^{XI}
	3.17 to 4.57	$w_1^I < w_1^{XI}$
	4.58 to 10	w_1^{XI}

Table B.6 (continued).

c	θ	Welfare Ranking
0.9	0.01 to 1.18	no interior solutions
	1.19	$w_4^{VIII} < w_2^{VIII} < w_1^{VIII}$
	1.2 to 1.28	$w_2^{V} < w_4^{VIII} < w_2^{VIII} < w_1^{VIII} < w_1^{V}$
	1.29 to 1.37	$w_2^{V} < w_4^{VIII} < w_3^{VII} < w_2^{VIII} < w_1^{XI} < w_1^{VII} < w_1^{VIII} < w_1^{V}$
	1.38 to 1.39	$w_2^{V} < w_4^{VIII} < w_3^{VII} < w_2^{VIII} < w_1^{IX} = w_1^{XI} < w_1^{VII} < w_1^{VIII} < w_1^{V}$
	1.4 to 1.48	$w_4^{VIII} < w_2^{V} < w_3^{VII} < w_2^{VIII} < w_1^{IX} = w_1^{XI} < w_1^{VII} < w_1^{VIII} < w_1^{V}$
	1.49 to 1.51	$w_2^{V} < w_3^{VII} < w_1^{IX} = w_1^{XI} < w_1^{VII} < w_1^{V}$
	1.52 to 1.58	$w_3^{VII} < w_1^{IX} = w_1^{XI} < w_1^{VII}$
	1.59 to 1.84	$w_1^{IX} = w_1^{XI}$
	1.85 to 1.94	$w_1^{X} < w_1^{IX} = w_1^{XI} < w_3^{X}$
	1.95 to 2.04	$w_1^{X} < w_1^{IV} < w_1^{IX} < w_1^{XI} < w_3^{X}$
	2.05 to 2.23	$w_1^{X} < w_1^{IV} < w_1^{XI} < w_3^{X}$
	2.24 to 2.77	$w_1^{X} < w_1^{XI} < w_3^{X}$
	2.78 to 3.4	w_1^{XI}
	3.41 to 4.6	$w_1^{I} < w_1^{XI}$
	4.61 to 10	w_1^{XI}

Table B.7 – Equilibria in the absence of FTAs ($n = 4$).

c	θ	NE	BS	GS	CPNE
0.1	0.01 to 1.72	n/a	n/a	n/a	n/a
	1.73 to 2.01	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI
	2.02 to 2.35	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI
	2.36 to 3.16	I, IV, V, VII, VIII, IX, X, XI	I, IV, V, VII, VIII, IX, X, XI	I, IV, V, VII, VIII, IX, X, XI	I, IV, V, VII, VIII, IX, X, XI
	3.17 to 4.09	I,IV	I,IV	IV	IV
	4.1 to 4.51	I	I	I	I
	4.52 to 10	n/a	n/a	n/a	n/a
0.3	0.01 to 1.87	n/a	n/a	n/a	n/a
	1.88 to 2.37	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI
	2.38 to 2.91	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI
	2.92 to 3.16	I, IV, V, VII, VIII, IX, X, XI	I, IV, V, VII, VIII, IX, X, XI	I, IV, V, VII, VIII, IX, X, XI	I, IV, V, VII, VIII, IX, X, XI
	3.17 to 4.1	I,IV	I,IV	IV	IV
	4.11 to 4.55	I	I	I	I
	4.56 to 10	n/a	n/a	n/a	n/a
0.5	0.01 to 1.93	n/a	n/a	n/a	n/a
	1.94 to 2.54	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI
	2.55 to 3.16	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI
	3.17 to 4.16	I,IV	I,IV	IV	IV
	4.17 to 4.57	I	I	I	I
	4.58 to 10	n/a	n/a	n/a	n/a

Table B.7 (continued).

c	θ	NE	BS	GS	CPNE
0.7	0.01 to 1.96	n/a	n/a	n/a	n/a
	1.97 to 2.64	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI
	2.65 to 3.15	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI
	3.16 to 3.3	IV	IV	IV	IV
	3.31 to 4.16	I,IV	I,IV	IV	IV
	4.17 to 4.59	I	I	I	I
	4.6 to 10	n/a	n/a	n/a	n/a
0.9	0.01 to 1.99	n/a	n/a	n/a	n/a
	2 to 2.71	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI	V, VII, VIII, IX, X, XI
	2.72 to 3.16	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI	IV, V, VII, VIII, IX, X, XI
	3.17 to 3.3	IV	IV	IV	IV
	3.31 to 4.16	I,IV	I,IV	IV	IV
	4.17 to 4.59	I	I	I	I
	4.6 to 10	n/a	n/a	n/a	n/a

Table B.8 – Equilibria in the presence of FTAs ($n = 4$).

c	θ	NE	BS	GS	CPNE
0.1	0.01 to 0.15	V,VII,VIII	VIII	VIII	VIII
	0.16 to 0.53	VII,VIII	VIII	VIII	VIII
	0.54 to 0.7	VIII	VIII	VIII	VIII
	0.71 to 1.24	n/a	n/a	n/a	n/a
	1.25 to 1.33	XI	XI	XI	XI
	1.34 to 1.41	IX,XI	IX,XI	IX,XI	IX,XI
	1.42	IX	IX	IX	IX
	1.43 to 2.04	IV,IX	IV,IX	IX	IV, IX
	2.05 to 2.09	IV,X	IV	none	IV
	2.1 to 2.35	X	X	X	X
	2.36 to 2.38	I,X	I	none	I
	2.39 to 2.44	I,II,X	II	none	II
	2.45 to 2.77	I,X	I	none	I
	2.78 to 4.51	I,XI	I,XI	XI	XI
	4.52 to 10	XI	XI	XI	XI
0.3	0.01 to 0.65	n/a	n/a	n/a	n/a
	0.66 to 0.77	V	V	V	V
	0.78 to 0.8	V,VII	VII	VII	VII
	0.81 to 1.05	V,VII,VIII	none	none	none
	1.06 to 1.16	VII,VIII	none	none	none
	1.17	VIII	VIII	VIII	VIII
	1.18 to 1.25	n/a	n/a	n/a	n/a
	1.26 to 1.35	XI	XI	XI	XI
	1.36 to 1.65	IX,XI	IX,XI	IX,XI	IX,XI
	1.66 to 1.73	cIX	IX	IX	IX
	1.74 to 2.04	IV,IX	IV,IX	IX	IV,IX
	2.05 to 2.16	IV,X	IV	none	IV
	2.17 to 2.77	X	X	X	X
	2.78 to 2.91	XI	XI	XI	XI
	2.92 to 4.55	I,XI	I,XI	XI	XI
	4.56 to 10	XI	XI	XI	XI
0.5	0.01 to 0.96	n/a	n/a	n/a	n/a
	0.97 to 1.02	V	V	V	V
	1.03 to 1.06	V,VIII	VIII	VIII	VIII
	1.07 to 1.24	V,VII,VIII	none	none	none
	1.25 to 1.27	V,VII,VIII,XI	XI	XI	XI
	1.28 to 1.3	V,VII,XI	VII,XI	XI	XI
	1.31 to 1.36	VII,XI	VII,XI	XI	XI
	1.37	VII,IX,XI	IX,XI	IX,XI	IX,XI
	1.38 to 1.75	IX,XI	IX,XI	IX,XI	IX,XI
	1.76 to 1.84	IX	IX	IX	IX
	1.85 to 2.04	IV,IX	IV,IX	IX	IV,IX
	2.05 to 2.19	IV,X	IV	none	IV
	2.2 to 2.77	X	X	X	X
	2.78 to 3.16	XI	XI	XI	XI
	3.17 to 4.57	I,XI	I,XI	XI	XI
	4.58 to 10	XI	XI	XI	XI

Table B.8 (continued).

c	θ	NE	BS	GS	CPNE
0.7	0.01 to 1.12	n/a	n/a	n/a	n/a
	1.13 to 1.12	V	V	V	V
	1.13 to 1.2	V,VIII	VIII	VIII	VIII
	1.21 to 1.25	V,VII,VIII	none	none	none
	1.26 to 1.37	V,VII,VIII,XI	XI	XI	XI
	1.38 to 1.43	V,VII,VIII,IX,XI	IX,XI	IX,XI	IX,XI
	1.44 to 1.5	VII,IX,XI	IX,XI	IX,XI	IX,XI
	1.51 to 1.8	IX,XI	IX,XI	IX,XI	IX,XI
	1.81 to 1.9	IX	IX	IX	IX
	1.91 to 2.04	IV,IX	IV,IX	IX	IV,IX
	2.05 to 2.21	IV,X	IV	none	IV
	2.22 to 2.77	X	X	X	X
	2.78 to 3.3	XI	XI	XI	XI
	3.31 to 4.59	I,XI	I,XI	XI	XI
	4.6 to 10	XI	XI	XI	XI
0.9	0.01 to 1.18	n/a	n/a	n/a	n/a
	1.19	VIII	VIII	VIII	VIII
	1.2 to 1.28	V,VIII	VIII	VIII	VIII
	1.29 to 1.37	V,VII,VIII,XI	XI	XI	XI
	1.38 to 1.39	V,VII,VIII,IX,XI	IX,XI	IX,XI	IX,XI
	1.4 to 1.48	V,VII,IX,XI	IX,XI	IX,XI	IX,XI
	1.49 to 1.51	V,VII,IX,XI	IX,XI	IX,XI	IX,XI
	1.52 to 1.58	VII,IX,XI	IX,XI	IX,XI	IX,XI
	1.59 to 1.84	IX,XI	IX,XI	IX,XI	IX,XI
	1.85 to 1.94	IX	IX	IX	IX
	1.95 to 2.04	IV,IX	IV,IX	IX	IV,IX
	2.05 to 2.23	IV,X	IV	none	IV
	2.24 to 2.77	X	X	X	X
	2.78 to 3.4	XI	XI	XI	XI
	3.41 to 4.6	I,XI	I,XI	XI	XI
	4.61 to 10	XI	XI	XI	XI

Table B.9 – Rankings of global welfare across trade regimes ($n = 3$).

c	θ	Global Welfare Ranking
0.1	0.01 to 1.33	no interior solutions
	1.34 to 1.82	gw^{III}
	1.83 to 1.92	no interior solutions
	1.93 to 2.43	gw^{IV}
	2.44 to 2.52	$gw^{II} < gw^{IV}$
	2.53 to 2.76	$gw^{II} < gw^{IV} = gw^{III-noFTA} < gw^{IV-noFTA}$
	2.77 to 3.1	$gw^{I} < gw^{II} < gw^{IV} = gw^{III-noFTA} < gw^{IV-noFTA}$
	3.11 to 4.08	$gw^{I} < gw^{IV} = gw^{III-noFTA} < gw^{IV-noFTA}$
	4.09 to 5.11	$gw^{I} < gw^{IV}$
	5.12 to 10	gw^{IV}
0.3	0.01 to 1.77	no interior solutions
	1.78 to 1.92	gw^{III}
	1.93 to 2.07	$gw^{III} < gw^{IV}$
	2.08 to 3.23	gw^{IV}
	3.24 to 3.43	$gw^{II} < gw^{IV} = gw^{III-noFTA} < gw^{IV-noFTA}$
	3.44 to 3.45	$gw^{I} < gw^{II} < gw^{IV} = gw^{III-noFTA} < gw^{IV-noFTA}$
	3.46 to 4.08	$gw^{I} < gw^{IV} = gw^{III-noFTA} < gw^{IV-noFTA}$
	4.09 to 5.17	$gw^{I} < gw^{IV}$
	5.18 to 10	gw^{IV}
0.5	0.01 to 1.91	no interior solutions
	1.92	gw^{III}
	1.93 to 2.17	$gw^{III} < gw^{IV}$
	2.18 to 3.73	gw^{IV}
	3.74 to 4.08	$gw^{I} < gw^{IV} = gw^{III-noFTA} < gw^{IV-noFTA}$
	4.09 to 5.21	$gw^{I} < gw^{IV}$
	5.22 to 10	gw^{IV}

Table B.9 (continued).

c	θ	Global Welfare Ranking
0.7	0.01 to 1.92	no interior solutions
	1.93 to 1.98	gw^{IV}
	1.99 to 2.22	$gw^{III} < gw^{IV}$
	2.23 to 3.91	gw^{IV}
	3.92 to 4.08	$gw^{I} < gw^{IV} = gw^{III-noFTA} < gw^{IV-noFTA}$
	4.09 to 5.24	$gw^{I} < gw^{IV}$
	5.25 to 10	gw^{IV}
0.9	0.01 to 1.92	no interior solutions
	1.93 to 2.02	gw^{IV}
	2.03 to 2.26	$gw^{III} < gw^{IV}$
2.27 to 4.03	gw^{IV}	
	4.04 to 4.08	$gw^{I} < gw^{IV} = gw^{III-noFTA} < gw^{IV-noFTA}$
	4.09 to 5.26	$gw^{I} < gw^{IV}$
	5.27 to 10	gw^{IV}

Table B.10 – Rankings of global welfare across trade regimes ($n = 4$).

c	θ	Global Welfare Ranking
0.1	0.01 to 0.15	$gw^{V} < gw^{VII} < gw^{VIII}$
	0.16 to 0.53	$gw^{VII} < gw^{VIII}$
	0.54 to 0.7	gw^{III}
	0.71 to 1.24	no interior solutions
	1.25 to 1.33	gw^{XI}
	1.34 to 1.41	$gw^{IX} = gw^{XI}$
	1.42	$gw^{X} < gw^{IX} = gw^{XI}$
	1.43 to 1.72	$gw^{X} < gw^{IV} < gw^{IX} = gw^{XI}$
	1.73 to 2.01	$gw^{X} < gw^{IV} < gw^{IX} = gw^{XI} = gw^{V-noFTA}$
	2.02 to 2.04	$gw^{X} < gw^{IV-noFTA} < gw^{IV} < gw^{IX} = gw^{XI} = gw^{V-noFTA}$
	2.05 to 2.09	$gw^{X} < gw^{IV-noFTA} < gw^{IV} < gw^{XI} = gw^{V-noFTA}$
	2.1 to 2.35	$gw^{IV-noFTA} < gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.36 to 2.38	$gw^{I} < gw^{IV-noFTA} < gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.39 to 2.44	$gw^{I} < gw^{IV-noFTA} < gw^{II} < gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.45 to 2.77	$gw^{I} < gw^{IV-noFTA} < gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.78 to 3.16	$gw^{I} < gw^{IV-noFTA} < gw^{XI} = gw^{V-noFTA}$
	3.17 to 4.09	$gw^{I} < gw^{IV-noFTA} < gw^{XI}$
	4.1 to 4.51	$gw^{I} < gw^{XI}$
	4.52 to 10	gw^{XI}

Table B.10 (continued).

c	θ	Global Welfare Ranking
0.3	0.01 to 0.65	no interior solutions
	0.66 to 0.77	gw^{V}
	0.78 to 0.8	$gw^{V} < gw^{VII}$
	0.81 to 1.05	$gw^{V} < gw^{VII} < gw^{VIII}$
	1.06 to 1.16	$gw^{VII} < gw^{VIII}$
	1.17	gw^{VIII}
	1.18 to 1.25	no interior solutions
	1.26 to 1.35	gw^{XI}
	1.36 to 1.65	$gw^{IX} = gw^{XI}$
	1.66 to 1.73	$gw^{X} < gw^{IX} = gw^{XI}$
	1.74 to 1.87	$gw^{X} < gw^{IV} < gw^{IX} = gw^{XI}$
	1.88 to 2.04	$gw^{X} < gw^{IV} < gw^{IX} = gw^{XI} = gw^{V-noFTA}$
	2.05 to 2.16	$gw^{X} < gw^{IV} < gw^{XI} = gw^{V-noFTA}$
	2.17 to 2.37	$gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.38 to 2.77	$gw^{IV-noFTA} < gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.78 to 2.91	$gw^{IV-noFTA} < gw^{XI} = gw^{V-noFTA}$
	2.92 to 3.16	$gw^{I} < gw^{IV-noFTA} < gw^{XI} = gw^{V-noFTA}$
	3.17 to 4.1	$gw^{I} < gw^{IV-noFTA} < gw^{XI}$
	4.11 to 4.55	$gw^{I} < gw^{XI}$
	4.56 to 10	gw^{XI}

Table B.10 (continued).

c	θ	Global Welfare Ranking
0.5	0.01 to 0.96	no interior solutions
	0.97 to 0.77	gw^{V}
	0.78 to 1.02	gw^{V}
	1.03 to 1.06	$gw^{V} < gw^{VIII}$
	1.07 to 1.24	$gw^{V} < gw^{VII} < gw^{VIII}$
	1.25 to 1.27	$gw^{V} < gw^{VII} < gw^{VIII} < gw^{XI}$
	1.28 to 1.3	$gw^{V} < gw^{VII} < gw^{XI}$
	1.31 to 1.36	$gw^{VII} < gw^{XI}$
	1.37	$gw^{VII} < gw^{IX} = gw^{XI}$
	1.38 to 1.75	$gw^{IX} = gw^{XI}$
	1.76 to 1.84	$gw^{X} < gw^{IX} = gw^{XI}$
	1.85 to 1.93	$gw^{X} < gw^{IX} = gw^{XI}$
	1.94 to 2.04	$gw^{X} < gw^{IV} < gw^{IX} = gw^{XI} = gw^{V-noFTA}$
	2.05 to 2.19	$gw^{X} < gw^{IV} < gw^{XI} = gw^{V-noFTA}$
	2.2 to 2.54	$gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.55 to 2.77	$gw^{IV-noFTA} < gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.78 to 3.16	$gw^{IV-noFTA} < gw^{XI} = gw^{V-noFTA}$
	3.17 to 4.16	$gw^{I} < gw^{IV-noFTA} < gw^{XI}$
	4.17 to 4.57	$gw^{I} < gw^{XI}$
	4.58 to 10	gw^{XI}

Table B.10 (continued).

c	θ	Global Welfare Ranking
0.7	0.01 to 1.1	no interior solutions
	1.11 to 1.12	gw^{V}
	1.13 to 1.2	$gw^{V} < gw^{VIII}$
	1.21 to 1.25	$gw^{V} < gw^{VII} < gw^{VIII}$
	1.26 to 1.37	$gw^{V} < gw^{VII} < gw^{VIII} < gw^{XI}$
	1.38 to 1.43	$gw^{V} < gw^{VII} < gw^{IX} = gw^{XI}$
	1.44 to 1.5	$gw^{VII} < gw^{IX} = gw^{XI}$
	1.51 to 1.8	$gw^{IX} = gw^{XI}$
	1.81 to 1.9	$gw^{X} < gw^{IX} = gw^{XI}$
	1.91 to 1.96	$gw^{X} < gw^{IV} < gw^{IX} = gw^{XI}$
	1.97 to 2.04	$gw^{X} < gw^{IV} < gw^{IX} = gw^{XI} = gw^{V-noFTA}$
	2.05 to 2.21	$gw^{X} < gw^{IV} < gw^{XI} = gw^{V-noFTA}$
	2.22 to 2.64	$gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.65 to 2.77	$gw^{IV-noFTA} < gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.78 to 3.15	$gw^{IV-noFTA} < gw^{XI} = gw^{V-noFTA}$
	3.16 to 3.3	$gw^{IV-noFTA} < gw^{XI}$
	3.31 to 4.17	$gw^{I} < gw^{IV-noFTA} < gw^{XI}$
	4.18 to 4.59	$gw^{I} < gw^{XI}$
	4.6 to 10	gw^{XI}

Table B.10 (continued).

c	θ	Global Welfare Ranking
0.9	0.01 to 1.18	no interior solutions
	1.19	gw^{VIII}
	1.2 to 1.24	$gw^{V} < gw^{VIII}$
	1.25 to 1.28	$gw^{V} < gw^{VII} < gw^{VIII}$
	1.29 to 1.37	$gw^{V} < gw^{VII} < gw^{VIII} < gw^{XI}$
	1.38 to 1.39	$gw^{V} < gw^{VII} < gw^{VIII} < gw^{IX} = gw^{XI}$
	1.4 to 1.51	$gw^{V} < gw^{VII} < gw^{IX} = gw^{XI}$
	1.52 to 1.58	$gw^{V} < gw^{IX} = gw^{XI}$
	1.59 to 1.84	$gw^{IX} = gw^{XI}$
	1.85 to 1.94	$gw^{X} < gw^{IX} = gw^{XI}$
	1.95 to 1.99	$gw^{X} < gw^{IV} < gw^{IX} = gw^{XI}$
	2 to 2.04	$gw^{X} < gw^{IV} < gw^{IX} = gw^{XI} = gw^{V-noFTA}$
	2.05 to 2.23	$gw^{X} < gw^{IV} < gw^{XI} = gw^{V-noFTA}$
	2.24 to 2.71	$gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.72 to 2.77	$gw^{IV-noFTA} < gw^{X} < gw^{XI} = gw^{V-noFTA}$
	2.78 to 3.16	$gw^{IV-noFTA} < gw^{XI} = gw^{V-noFTA}$
	3.17 to 3.4	$gw^{IV-noFTA} < gw^{XI}$
	3.41 to 4.17	$gw^{I} < gw^{IV-noFTA} < gw^{XI}$
	4.18 to 4.6	$gw^{I} < gw^{XI}$
	4.61 to 10	gw^{XI}

REFERENCES

[1] N. Allouch, "On the private provision of public goods on networks," *Journal of Economic Theory*, vol. 157, pp. 527–552, 2015.

[2] J. Andreoni, "Giving with impure altruism: Applications to charity and ricardian equivalence," *Journal of Political Economy*, vol. 97, no. 6, pp. 1447–1458, 1989.

[3] ——, "Impure altruism and donations to public goods: A theory of warm-glow giving," *The Economic Journal*, vol. 100, no. 401, pp. 464–477, 1990.

[4] ——, "Giving gifts to groups: How altruism depends on the number of recipients," *Journal of Public Economics*, vol. 91, no. 9, pp. 1731–1749, 2007.

[5] A. Badev, "Discrete games in endogenous networks: Theory and policy," *Working Paper*, 2016.

[6] V. Bala and S. Goyal, "A noncooperative model of network formation," *Econometrica*, vol. 68, no. 5, pp. 1181–1229, 2000.

[7] S. Barrett, "Self-enforcing international environmental agreements," *Oxford Economic Papers*, pp. 878–894, 1994.

[8] ——, "The incredible economics of geoengineering," *Environmental and resource economics*, vol. 39, no. 1, pp. 45–54, 2008.

[9] M. Belhaj, Y. Bramoullé, and F. Deroïan, "Network games under strategic complementarities," *Games and Economic Behavior*, vol. 88, pp. 310–319, 2014.

[10] R. Bénabou and J. Tirole, "Incentives and prosocial behavior," *American Economic Review*, vol. 96, no. 5, pp. 1652–1678, 2006.

[11] A. BenYishay and A. M. Mobarak, "Social learning and communication," *NBER Working Paper*, 2014.

[12] B. D. Bernheim, B. Peleg, and M. D. Whinston, "Coalition-proof nash equilibria i. concepts," *Journal of Economic Theory*, vol. 42, no. 1, pp. 1–12, 1987.

[13] R. Bourlès, Y. Bramoullé, and E. Perez-Richet, "Altruism in networks," *Econometrica*, vol. 85, no. 2, pp. 675–689, 2017.

[14] Y. Bramoullé and R. Kranton, “Public goods in networks,” *Journal of Economic Theory*, vol. 135, no. 1, pp. 478–494, 2007.

[15] W. Buchholz, R. Cornes, and D. Rübbelke, “Interior matching equilibria in a public good economy: An aggregative game approach,” *Journal of Public Economics*, vol. 95, no. 7-8, pp. 639–645, 2011.

[16] A. J. Caplan and E. C. Silva, “An efficient mechanism to control correlated externalities: Redistributive transfers and the coexistence of regional and global pollution permit markets,” *Journal of Environmental Economics and Management*, vol. 49, no. 1, pp. 68–82, 2005.

[17] C. Carraro and D. Siniscalco, “Strategies for the international protection of the environment,” *Journal of Public Economics*, vol. 52, no. 3, pp. 309–328, 1993.

[18] R. Cornes and R. Hartley, “Aggregative public good games,” *Journal of Public Economic Theory*, vol. 9, no. 2, pp. 201–219, 2007.

[19] R. Cornes and T. Sandler, “Easy riders, joint production, and public goods,” *The Economic Journal*, vol. 94, no. 375, pp. 580–598, 1984.

[20] ——, “Easy riders, joint production, and public goods,” *The Economic Journal*, vol. 94, no. 375, pp. 580–598, 1984.

[21] E. Diamantoudi and E. S. Sartzetakis, “Stable international environmental agreements: An analytical approach,” *Journal of Public Economic Theory*, vol. 8, no. 2, pp. 247–263, 2006.

[22] E. Diamantoudi, E. S. Sartzetakis, and S. Strantza, “International environmental agreements and trading blocks - can issue linkage enhance cooperation?” *FEEM Working Paper*, 2018.

[23] M. Dufwenberg and A. Patel, “Reciprocity networks and the participation problem,” *Games and Economic Behavior*, vol. 101, pp. 260–272, 2017.

[24] T. Eichner and R. Pethig, “Self-enforcing environmental agreements and international trade,” *Journal of Public Economics*, vol. 102, pp. 37–50, 2013.

[25] ——, “Is trade liberalization conducive to the formation of climate coalitions?” *International Tax and Public Finance*, vol. 22, no. 6, pp. 932–955, 2015.

[26] ——, “Self-enforcing international environmental agreements and trade: Taxes versus caps,” *Oxford Economic Papers*, vol. 67, no. 4, pp. 897–917, 2015.

[27] S. Erol and R. Vohra, “Network formation and systemic risk,” *Working Paper*, 2018.

[28] L. F. Feick and L. L. Price, “The market maven: A diffuser of marketplace information,” *The Journal of Marketing*, pp. 83–97, 1987.

[29] D. Foster and P. Young, “Stochastic evolutionary game dynamics,” *Theoretical Population Biology*, vol. 38, no. 2, p. 219, 1990.

[30] A. Galeotti, “One-way flow networks: The role of heterogeneity,” *Economic Theory*, vol. 29, no. 1, pp. 163–179, 2006.

[31] A. Galeotti and S. Goyal, “The law of the few,” *American Economic Review*, vol. 100, no. 4, pp. 1468–92, 2010.

[32] A. Galeotti, S. Goyal, and J. Kamphorst, “Network formation with heterogeneous players,” *Games and Economic Behavior*, vol. 54, no. 2, pp. 353–372, 2006.

[33] T. Goeschl, D. Heyen, and J. Moreno-Cruz, “The intergenerational transfer of solar radiation management capabilities and atmospheric carbon stocks,” *Environmental and resource economics*, vol. 56, no. 1, pp. 85–104, 2013.

[34] T. Hiller, “Peer effects in endogenous networks,” *Games and Economic Behavior*, vol. 105, pp. 349–367, 2017.

[35] N. Immorlica, R. Kranton, M. Manea, and G. Stoddard, “Social status in networks,” *American Economic Journal: Microeconomics*, vol. 9, no. 1, pp. 1–30, 2017.

[36] M. Kandori, G. J. Mailath, and R. Rob, “Learning, mutation, and long run equilibria in games,” *Econometrica*, pp. 29–56, 1993.

[37] M. Kinateder and L. P. Merlino, “Public goods in endogenous networks,” *American Economic Journal: Microeconomics*, vol. 9, no. 3, pp. 187–212, 2017.

[38] T. Kuhn, R. Pestow, and A. Zenker, “Building climate coalitions on preferential free trade agreements,” *Environmental and Resource Economics*, vol. 74, no. 2, pp. 539–569, 2019.

[39] B. van Leeuwen, T. Offerman, and A. J. Schram, “Competition for status creates superstars: An experiment on public good provision and network formation,” *Working Paper*, 2015.

[40] J. Lerner and J. Tirole, “The economics of technology sharing: Open source and beyond,” *Journal of Economic Perspectives*, vol. 19, no. 2, pp. 99–120, 2005.

[41] A. Millard-Ball, “The tuvalu syndrome,” *Climatic Change*, vol. 110, no. 3-4, pp. 1047–1066, 2012.

[42] J. B. Moreno-Cruz, “Mitigation and the geoengineering threat,” *Working Paper*, 2011.

[43] J. B. Moreno-Cruz and D. W. Keith, “Climate policy under uncertainty: A case for solar geoengineering,” *Climatic Change*, vol. 121, no. 3, pp. 431–444, 2013.

[44] J. B. Moreno-Cruz, K. L. Ricke, and D. W. Keith, “A simple model to account for regional inequalities in the effectiveness of solar radiation management,” *Climatic change*, vol. 110, no. 3-4, pp. 649–668, 2012.

[45] W. Nordhaus, “Climate clubs: Overcoming free-riding in international climate policy,” *American Economic Review*, vol. 105, no. 4, pp. 1339–70, 2015.

[46] L. Richefort, “Warm-glow giving in networks with multiple public goods,” *International Journal of Game Theory*, pp. 1–28, 2017.

[47] S. J. Rubio and A. Ulph, “Self-enforcing international environmental agreements revisited,” *Oxford Economic Papers*, vol. 58, no. 2, pp. 233–263, 2006.

[48] E. C. D. Silva and C. Yamaguchi, “Overlapping climate clubs under transaction costs,” *CESifo Working Paper*, 2018.

[49] E. C. D. Silva, X. Zhu, *et al.*, “Overlapping international environmental agreements,” *Strategic Behavior and the Environment*, vol. 5, no. 3–4, pp. 255–299, 2015.

[50] J. Urpelainen, “Geoengineering and global warming: A strategic perspective,” *International Environmental Agreements: Politics, Law and Economics*, vol. 12, no. 4, pp. 375–389, 2012.

[51] H. R. Varian, *Microeconomic analysis*. Norton New York, 1992, vol. 3.

[52] V. V. Vasconcelos, P. M. Hannam, S. A. Levin, and J. M. Pacheco, “Coalition-structured governance improves cooperation to provide public goods,” *arXiv preprint arXiv:1910.11337*, 2019.

[53] G. Walsh, K. P. Gwinner, and S. R. Swanson, “What makes mavens tick? exploring the motives of market mavens’ initiation of information diffusion,” *Journal of Consumer Marketing*, vol. 21, no. 2, pp. 109–122, 2004.

[54] H. P. Young, “The evolution of conventions,” *Econometrica*, pp. 57–84, 1993.

[55] X. M. Zhang and F. Zhu, “Group size and incentives to contribute: A natural experiment at chinese wikipedia,” *American Economic Review*, vol. 101, no. 4, pp. 1601–15, 2011.