Analysing International Environmental Agreements: Approaches and Equilibrium Concepts

Michael Finus

University of Bath, Department of Economics, UK

Based on Finus (2001) as wells as

(2003) and (2008) overview articles plus further literature.

A Literature Guide will be prepared and provided next week.

- externalities, (pure and impure) public goods, common pool resources
- international, global, transboundary
- cooperation (formal or informal agreements), IEAs
- game theory (cooperative, non-cooperative)
- Barrett (1994), Carraro and Siniscalco (1993), Chander and Tulkens (1992) and Hoel (1992)
- How can we capture and explain the free-riding problems in an international environmental externality game? (positive analysis)
- How can we mitigate the free-riding problems? (normative analysis)

Structure of IEA Models

Main Features	Sub-Features	Characteristics	
Time	framework	implicit dynamic	explicit dynamic
	horizon		finite or infinite
	interval		discrete or continuous
Payoff	structural relation	independent (flow pollution)	dependent (stock pollution)
	arguments	only material payoffs	also non-material payoffs
	transfers	no	yes

Main Features	Sub-Features	Characteristics	
equilibrium concepts	strategic relation	independent	dependent
	sanctions	different degrees of harshness and credibility of sanctions	
	deviations	single	multiple
number of issues		single	multiple
rules of coalition formation	sequence of coalition formation	simultaneous	sequential
	number of coalitions	single	multiple
	membership	open	exclusive
	consensus	different degrees of consensus with respect to membership	

IEA

membership models
(free-riding of type 1)

cooperative vs non-cooperative
coalition models

compliance models
(free-riding of type 2)

non-cooperative
repeated game

a) Difference between membership and compliance models?

b) Difference between cooperative and non-cooperative coalition models?

Membership Models

Static Social Dilemma Game

		2			
		C		D	
1	C	2	2	-1	3
	D	3	-1	0	0

Assumption:

Full cooperation pays globally: 4>0 and 4>-1+3=2

Full cooperation pays individually: 2>0

Individual cooperation does not pay: 0<-1

Assumption:

Cooperation of n countries generates $n \bullet B$ benefits.

Cooperation implies C cost.

Hence if two countries cooperate, their payoff is given by $\Pi = 2B - C$.

If one country cooperates and the other free-rides, then the co-operator receives $\Pi = B - C$ and the free-rider $\Pi = B$.

If both free-ride, then $\Pi = 0$.

Full cooperation pays globally: $2(2B - C) > 2B - C$, $2(2B - C) > 0 \Leftrightarrow 2B > C$

Full cooperation pays individually: $2B - C > 0 \Leftrightarrow 2B > C$.

Individual cooperation does not pay: $B - C < 0 \Leftrightarrow B < C$

Example: $B = 3$ and $C = 4$.

		2			
		C		D	
1	C	2	2	-1	3
	D	3	-1	0	0

Equilibrium in Dominant Strategies

Each player plays his/her best strategy, irrespective of the strategies of the other players.

Nash Equilibrium

A Nash equilibrium is a strategy combination $s^* = (s_i^*, s_{-i}^*)$ for which $\Pi_i(s_i^*, s_{-i}^*) \geq \Pi_i(s_i, s_{-i}^*)$ for all $s_i \neq s_i^*$ and all i. That is, in equilibrium, every strategy is a best response to the best strategy of all other players. Hence, strategies are mutual best replies.

What are the crucial assumptions for the negative result?

- discrete strategies ?
- payoff structure ?
- static game ?

Which other assumptions are not in line with reality?

- only two players ?
- symmetry ?

Static Chicken Game

		2			
		C		D	
1	C	2	2	0.5	3
	D	3	0.5	0	0

Assumption:

Full cooperation pays globally: 4>0 and 4>0.5+3=3.5

Full cooperation pays individually: 2>0

Individual cooperation does pay: 0.5<0

What are the crucial assumptions for the negative result?

- discrete strategies ?
- payoff structure ?
- static game ?

Which other assumptions are not in line with reality?

- only two players ?
- symmetry ?

Internal & External Stability in the Cartel Formation Game

Suppose there are N players.

1. Stage: Players decide whether to participate in an agreement.

2. Stage: Players choose their economic strategies.

a) Signatories cooperate (i.e. maximize their aggregate payoff).

b) Non-signatories do not cooperate (i.e. maximize their own payoff).

Stable Agreement

An agreement is stable if no signatory has an incentive to leave the agreement to become a non-signatory, and no non-signatory has an incentive to join the agreement.

Internal Stability: $\Pi_i(S) \geq \Pi_i(S \setminus \{i\})$ for all $i \in S$

External Stability: $\Pi_j(S) \geq \Pi_j(S \cup \{j\})$ for all $j \notin S$

Payoff Functions

Emission Space (e_i) or Abatement Space (q_i); static, pure global public bad/good

[1] $\pi_i = B_i(e_i) - D_i(\sum_{j=1}^{N} e_j)$

[2] $\pi_i = B_i(\sum_{j=1}^{N} q_j) - C_i(q_i)$

[3] $\pi_i = -C_i(e_i^0 - e_i) - D_i(\sum_{j=1}^{N} e_j)$

For [1] we would assume: $e_i \in [0, e_i^{max})$: $B_i' > 0$, $B_i'' < 0$, $D_i' > 0$, $D_i'' \geq 0$ $\forall$ i.

Linear-linear Payoff Function: “Toy Model”

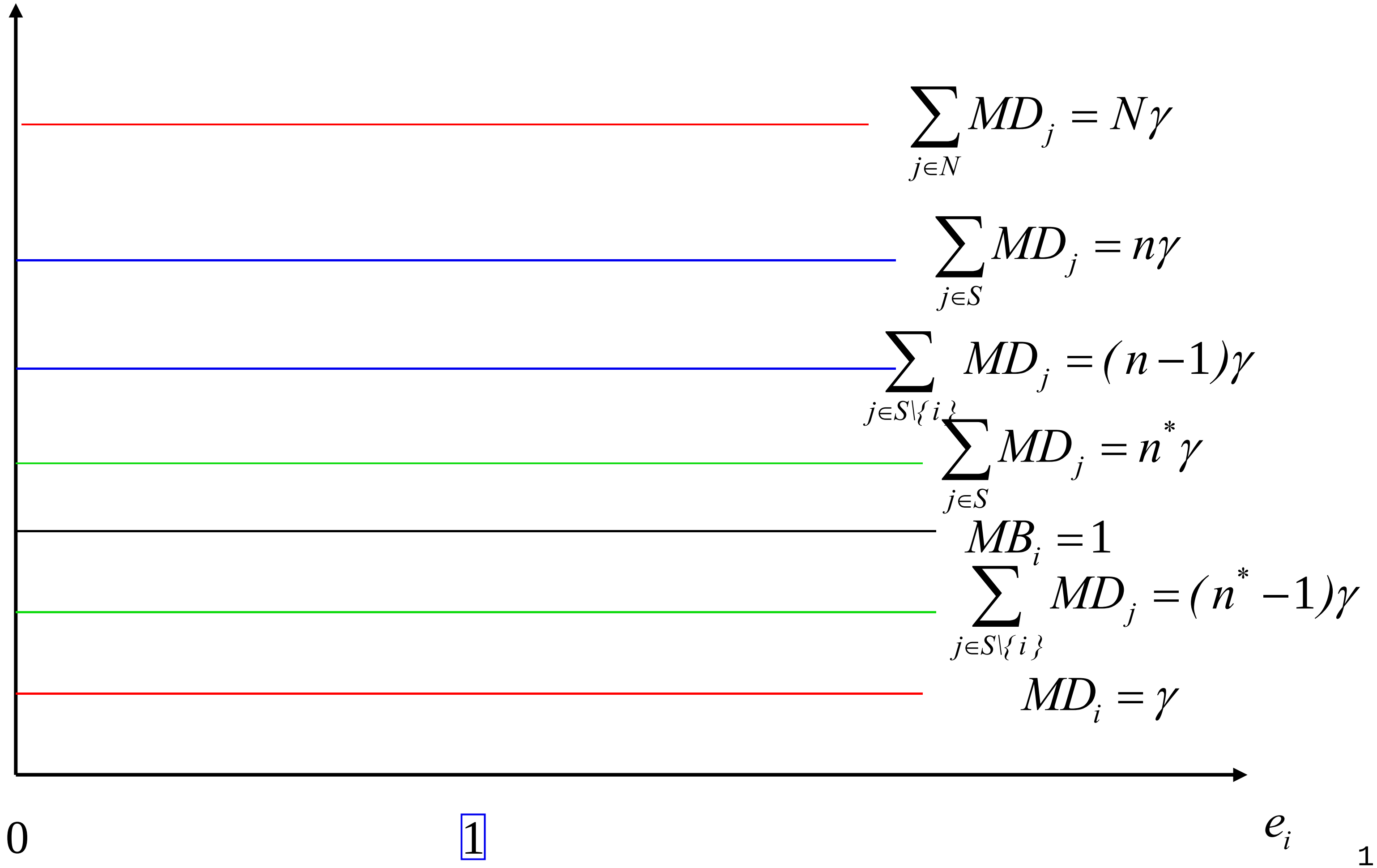
$\Pi_i(e_i, E) \equiv e_i - \gamma E$ with $E = \sum_{j=1}^{N} e_j$.

Strategies: $e_i = 1$ (pollute) or $e_i = 0$ (abate)

Assumption: $\frac{1}{N} < \gamma < 1$

Equilibrium Analysis

$$n^* = \text{integer of } \frac{1}{\gamma}$$



Equilibrium Coalition Size*

Payoff Functions		Slopes of the Best Reply Functions	Choice of Emissions	
			Nash-Cournot	Stackelberg
Type 1	$\pi_i = be_i - c(\sum_{k=1}^{N} e_k)$	0	$n^*=\text{int}[\gamma]$	$n^*=\text{int}[\gamma]$
Type 2	$\pi_i = b(ae_i - \frac{1}{2}e_i^2) - c(\sum_{k=1}^{N} e_k)$	0	$n^*=3$	$n^*=3$
Type 3	$\pi_i = b(ae_i - \frac{1}{2}e_i^2) - \frac{c}{2}(\sum_{k=1}^{N} e_k)^2$	between 0 and -1	$n^* \in \{1, 2\}^{\#}$	$n^* \in [2, N]^{\#}$, $\partial n^*/\partial\gamma<0$
Type 4	$\pi_i = be_i - \frac{c}{2}(\sum_{k=1}^{N} e_k)^2$	-1	$n^*=1$	$n^*=N$

* a, b, c, and N are parameters, where N denotes the total number of countries and γ=b/c; n* denotes the equilibrium number of participants; # means that the coalition size depends on parameter values. The results in the column "Choice of Emissions" apply to the examples.

What are the crucial assumptions of I&E-Stability if the Cartel Formation Game?

- implicit dynamic story
- single deviation
- remaining coalition members continue to cooperate after deviation
- single agreement
- open membership

Core-Stable Coalitions

Definition: Characteristic Function

Let I^J denote a subset of players forming a coalition and let I^{NJ} denote the set of all other players, then the worth of coalition I^J is given by $w(I^J) = \sum_{i \in I^J} \pi_i(e^J, e^{NJ})$ where emission vector e^{NJ} follows from some assumption about the behavior of players outside coalition I^J, I^{NJ}, and emission vector e^J from the maximization of the aggregate payoff of the players belonging to coalition I^J, $\max_{e^J} \sum_{i \in I^J} \pi_i(e^J, e^{NJ})$.

Definition: Core

An imputation $\pi^* = (\pi_1^*, \ldots, \pi_N^*)$ *lies in the core if* $\sum_{i \in I^J} \pi_i^* \geq w(I^J)\ \ \forall\, I^J \subset I$.

- term "coalition" is *not* used for the group of cooperating players but for the group of deviating players
- multiple deviations are possible
- whether deviations of coalition I^J are profitable depends on the assumption how the set of remaining players, I^{NJ}, will react.

- α-characteristic function, $w^{\alpha}(I^J)$, assumes that each player in I^{NJ} choose its highest emission level, e_j^{max}, in order to *minimize* the aggregate payoff of the deviators I^J.
- γ-characteristic function, $w^{\gamma}(I^J)$, assumes that after a deviation, the remaining countries break up into singletons, where each singleton maximizes its individual payoff in a Nash fashion.
- Because $w^{\gamma}(I^J) \geq w^{\alpha}(I^J)$, $C^{\gamma} \subset C^{\alpha}$ holds.

Result

- *only the grand coalition in which all players implement the socially optimal emission vector qualifies for being core-stable.*
- *Chander/Tulkens (1995 and 1997) establish that an imputation, which is derived from the socially optimal emission vector and where a particular transfer rule is applied, lies in the core.*
- *Helm (2001) establish generally that the core is non-empty in the global emission game based on some general properties.*

$$\sum_{i=1}^{N} t_i = 0 \text{ and } \pi_i^* = \pi_i^S + t_i$$

$$t_i = [\pi_i^N - \pi_i^S(e_i^S)] + \frac{D_i'(e^S)}{\sum_{k=1}^{N} D_k'(e^S)} \cdot [\sum_{k=1}^{N} \pi_k^S - \sum_{k=1}^{N} \pi_k^S]$$

Table: Historical Development of Models Applying the Concept of the Core

Paper	Payoff Structure					Pollution		Core
	static	dynamic						
		discret	cont.	flow	stock	global	transb.	
1) Tulkens (1979)			•	•			•	---
2) Chander/Tulkens (1991)			•	•			•	α; d, t
3) Chander/Tulkens (1992)			•	•			•	α, d, t
4) Kaitala/Mäler/Tulkens (1995)			•	•			•	α, d, t
5) Chander/Tulkens (1995)	•					•		γ, s, t
6) Chander/Tulkens (1997)	•					•		γ, s, t
7) Germain/Toint/Tulkens (1996a)		•		•			•	γ, d, t

Column "Core": d=dynamic stability test along the entire time path, s=static stability test, t=core properties theoretically established, e=core properties empirically established, c=closed loop strategies, o=open loop strategies.

Table: Historical Development of Models Applying the Concept of the Core

Paper	Payoff Structure					Pollution		Core
	static	dynamic						
		discret	cont.	flow	stock	global	transb.	
8) Germain/Toint/Tulkens (1996b)		●		●			●	γ, d, t
9) Germain/Toint/Tulkens (1998)			●		●	●		γ, s, t, o
10) Germain/Tulkens/de Zeeuw (1998)			●		●	●		γ, d, t, o
11) Germain/Toint/Tulkens, de Zeeuw (2000)			●		●	●		γ, d, t, c
12) Eyckmans/Tulkens (1999)			●		●	●		γ, s, e, o
13) Germain/ van Ypersele (1999)			●		●	●		γ, d, e, c

Column "Core": d=dynamic stability test along the entire time path, s=static stability test, t=core properties theoretically established, e=core properties empirically established, c=closed loop strategies, o=open loop strategies. References listed in Finus (2003).

What are the crucial assumptions of Core-Stability?

- implicit dynamic story
- *multiple deviation*
- *after a deviation, remaining coalition members break up*
- *grand coalition*
- *exclusive membership*

Is the optimistic view better than the pessimistic view?

Evaluation

- positive versus normative approach
- reactions after a deviation should follow from the rules of coalition formation
- non-cooperative versus cooperative approach
- rules of coalition formation should be separate from stability concept used to determine stable coalitions

Overview of “New Approach”

1) Membership Game with Membership Rules

2) Economic Game with Economic Rules

First Stage

Let the set of agents be given by I={1, …, N} with a particular agent denoted by i, j or k. In the first stage of coalition formation, agents announce their membership strategies. The decision depends on the membership rule that is captured by the definition of a coalition game. The definition comprises two elements:

- the Cartesian product of strategy spaces $\Sigma = \Sigma_1 \times \Sigma_2 \times ... \times \Sigma_N$ with a particular strategy of agent $i \in I$ denoted by $\sigma_i \in \Sigma_i$, and
- the coalition function ψ that maps membership strategies $\sigma = (\sigma_1, ..., \sigma_N)$ into coalition structures, $\psi : \Sigma \to C$, $\sigma \mapsto \psi(\sigma)$.

A coalition structure $c = \{c^1, ..., c^M\}$ is a partition of the set of agents. A particular coalition is denoted by c^ℓ, c^m, c^n or $c^{\ell(i)}$ where we use the last notation to indicate that the coalition contains a particular agent i. Thus, we have $c^\ell \cap c^m = \varnothing \ \forall \ \ell \neq m$, $\bigcup c^\ell = I$ and $c \in C$ where C is the set of all possible coalition structures.

Second Stage

In the second stage of coalition formation, agents choose their economic strategies. The decision depends on the economic rule that is captured by the definition of the partition function. The definition comprises three items: a) economic strategies $s_i \in S_i$, $s \in S = S_1 \times ... \times S_N$, b) a utility function $u: S \to U = U_1 \times ... \times U_N \subseteq \mathbb{R}^{\mathbf{N}}$, $s \mapsto u(s)$ and c) an instruction ε how agents choose their economic strategies for a given coalition structure c, $\varepsilon : C \to S$, $c \mapsto \varepsilon(c)$.

That is, the partition function is a composition of two functions $v = u \circ \varepsilon$ where ε is a function mapping coalition structures into a vector of economic strategies and u is a function mapping economic strategies into utility levels. The game is solved backward.

Definition 1: Open Membership Single Coalition Game (OMSCG)

a) The set of membership strategies of agent $i \in I$ *is given by* $\Sigma_i = \{0, 1, \}$ *where a particular strategy* σ_i *is an announcement of an address. All agents announce their strategies simultaneously.*

b) Coalition function ψ^{OMSCG} *maps strategy vector* σ *into coalition structure c as follows:*

$$c^{\ell(i)} = \begin{cases} \{i\} & \text{if } \sigma_i = 0 \\ \{i\} \cup \{ j \mid \sigma_j = \sigma_i = 1\} & \text{otherwise.} \end{cases}$$

Definition 2: Open Membership Multiple Coalition Game (OMMCG)

a) The set of membership strategies of agent $i \in I$ is given by $\Sigma_i = \{0, 1, \ldots, N\}$ where a particular strategy σ_i is an announcement of an address. All agents announce their strategies simultaneously.

b) Coalition function ψ^{OMG} maps strategy vector σ into coalition structure c as follows:

$$c^{\ell(i)} = \begin{cases} \{i\} & \text{if } \sigma_i = 0 \\ \{i\} \cup \{ j \mid \sigma_j = \sigma_i \} & \text{otherwise.} \end{cases}$$

Definition 3: Exclusive Membership Δ- and Γ-Game (EMΔG and EMΓG)

a) The set of membership strategies of agent $i \in I$ *is given by* $\Sigma_i = \{ c^{\ell} \subset I \mid i \in c^{\ell} \}$ *where a particular strategy* σ_i *is a list of agents with whom agent i would like to form a coalition. All agents announce their strategies simultaneously.*

b)

i) In the EMΔG, coalition function $\psi^{EM\Delta G}$ *maps strategy vector* σ *into coalition structure c as follows:* $c^{\ell(i)} = \{ i \} \cup \{ j \mid \sigma_i = \sigma_j \}$.

ii) In the EMΓG, coalition function $\psi^{EM\Gamma G}$ *maps strategy vector* σ *into coalition structure c as follows:* $c^{\ell(i)} = \sigma_i$ *if and only if* $\sigma_i = \sigma_j \;\; \forall j \in \sigma_i$, *otherwise* $c^{\ell(i)} = \{ i \}$.

….. Exclusive Membership J-, H- and I-Game (EMJG, EMHG and EMIG)

Finus and Rundshagen (2008).

Definition 5: Nash and Strong Nash Equilibrium Coalition Structures

Let $\hat{C}^{I^S}(\tilde{\sigma})$ *be the set of coalition structures that a subgroup of agents* I^S *can induce if the remaining agents* $I \backslash I^S$ *play* $\tilde{\sigma}^{I \backslash I^S}$. *That is,* $\hat{C}^{I^S}(\tilde{\sigma}) = \{ c(\sigma) | \sigma_i \in \Sigma_i \ \forall i \in I^S, \ \sigma_j = \tilde{\sigma}_j \ \forall j \in I \backslash I^S \}$. *Then* σ^*, *inducing coalition structure* $c^* = c(\sigma^*)$, *is called a SNE if no subgroup* I^S *can increase its members′ utility by inducing another coalition structure* $\hat{c} \in \hat{C}^{I^S}(\sigma^*)$. *That is,* $c^*(\sigma^*)$ *is a SNE if there is no* $I^S \subset I$ *such that there exists a coalition structure* $\hat{c} \in \hat{C}^{I^S}(\sigma^*)$ *such that* $v_i(\hat{c}) \geq v_i(c^*) \ \forall i \in I^S \wedge \exists j \in I^S\!: v_j(\hat{c}) > v_j(c^*)$. *For a NE,* $I^S = \{i\}$.

….. Coalition-Proof Nash Equilibrium

Bernheim/Whinston/Peleg (1987) applied in Finus and Rundshagen (2003).

Possible Analyses

1) Assume some general properties about the partition function (second stage) and concentrate on the first stage. Derive some general properties of stable coalition equilibria. Maybe show which economic problems lead to these properties.

2) Assume a particular economic problem and derive some more specific properties about the partition function (second stage). Derive some more specific properties of stable coalition equilibria.

Example 1: Finus and Rundshagen (2008)

Definition: Positive Externality Property from a Merger of Coalitions

Let a coalition structure with M coalitions be denoted by c and a coalition structure with M-1 coalitions by $\hat{c} \in \hat{C}(c)$ *where* $\hat{C}(c)$ *denotes the set of coalition structures* $\hat{c}$ *which may be derived from c by merging two coalitions* c^{ℓ} *and* c^{m} *in c, i.e.* $\exists\, c^{\ell},\ c^{m} \in c : \hat{c} = (c \backslash \{c^{\ell}, c^{m}\}) \cup \{c^{\ell} \cup c^{m}\}$. *Moreover, let* c^{n} *be a coalition not involved in the merger, i.e.* $c^{n} \in c \cap \hat{c}$. *Then valuations are characterized by positive externalities if* $\forall\, c \in C$, $\hat{c} \in \hat{C}(c)$, $c^{n} \in c \cap \hat{c}$ *and* $\forall\, k \in c^{n}$: $v_{k}(\hat{c}) \geq v_{k}(c)$.

Typical Problems with Positive Externalities

- public good problems
- Cournot-Oligopoly
- Cooperation on R&D
- …… see Bloch (2003) and Yi (1997, 2003)

Corollary: Comparison of Equilibria in All Coalition Games

Let $C^{NE}(...)$ and $C^{SNE}(...)$ denote the set of Nash equilibrium (NE) and strong Nash equilibrium (SNE) coalition structures in the open membership game (OMG) and the exclusive membership Δ-, Γ-, J-, H-, and I-game (EMΔG, EMΓG, EMJG, EMHG and EMIG), respectively, then in positive externality games

a) $C^{NE}(OMG) \subset C^{NE}(EM\Delta G) = C^{NE}(EMJG) \subset C^{NE}(EM\Gamma G) \subset C^{NE}(EMHG) = C^{NE}(EMIG)$ *and*

b) $C^{SNE}(OMG) \subset C^{SNE}(EM\Delta G) = C^{SNE}(EMJG) \subset C^{SNE}(EM\Gamma G) \subset C^{SNE}(EMHG) \subset C^{NE}(EMIG)$.

Example 2: Finus and Rundshagen (2003)

Look at specific functions (and assume symmetry)

[2] $$\pi_i = b(ae_i - \frac{1}{2}e_i^2) - d(\sum_{j=1}^{N} e_j)$$

[3] $$\pi_i = b\,ln(1+e_i) - d(\sum_{j=1}^{N} e_j)$$

[4] $$\pi_i = b(ae_i - \frac{1}{2}e_i^2) - \frac{d}{2}(\sum_{j=1}^{N} e_j)^2$$

[5] $$\pi_i = b\,ln(1+e_i) - \frac{d}{2}(\sum_{j=1}^{N} e_j)^2$$

and derive some specific properties …and then characterize coalition structures.

Example 3: Finus (2008) summary of a couple of papers

payoff functions derived from a calibrated climate model with 12 world regions

Overview of Possible Assumptions

1. Stage: participation strategies

sequence	simultaneous	sequential
agreements	single	multiple
membership	open	exclusive (majority, unanimity)
min. participation clause	no	yes

2. Stage: economic strategies

abatement	efficient	bargaining (majority, unanimity, modesty)
transfers	no	yes
payoff structure	static vs dynamic	continuous vs thresholds
parameter values	known	unkown
other strategies	R&D, trade measures, ..., …., adaptation, geoengineering	
other payoff components	other regarding preferences, ancillary benefits, ….	

Understanding the Incentive Structure

Definition: Positive Externality Games

Let a coalition structure with M coalitions be denoted by $c=\{c_1, ..., c_M\}$, a coalition structure with M-1 coalitions by $c' = \{c'_1, ..., c'_{M-1}\}$ where c' is derived by merging two coalitions c_ℓ and c_m in c, and let player k belong to coalition c_n that is not involved in the merger, then in positive externality games for all $k \in c_n$: $\pi_k(c) < \pi_k(c')$.

Definition: Superadditive Games

Let a coalition structure with M coalitions be denoted by $c=\{c_1, ..., c_M\}$, a coalition structure with M-1 coalitions by $c' = \{c'_1, ..., c'_{M-1}\}$ where c' is derived by merging two coalitions c_ℓ and c_m in c, then in a superadditive game: $\sum_{i \in c_\ell} \pi_i(c) + \sum_{j \in c_m} \pi_j(c) < \sum_{k \in c_\ell \cup c_m} \pi_k(c')$.

Definition: Global Efficiency from Cooperation

Let a coalition structure with M coalitions be denoted by $c=\{c_1, ..., c_M\}$, a coalition structure with M-1 coalitions by $c' = \{c'_1, ..., c'_{M-1}\}$ where c' is derived by merging two coalitions c_ℓ and c_m in c, then in a game with global efficiency of cooperation: $\sum_{i=1}^{N} \pi_i(c) < \sum_{i=1}^{N} \pi_i(c')$.

Definition: Global Optimality

Let a coalition structure with M coalitions be denoted by $c=\{c_1, ..., c_M\}$, a coalition structure with one coalition (i.e. the grand coalition) by c^G, then the game is globally optimal if for all coalition structures $c \neq c^G$: $\sum_{i=1}^{N} \pi_i(c) < \sum_{i=1}^{N} \pi_i(c^G)$.

How can we mitigate free-riding?

Strengthen superadditivity effect?

Economies of scale superadditivity?

Reduce positive externality effect?

Turn game into negative externality game?