Ecological Economics 85 (2013) 177–187

Contents lists available at ScienceDirect

Ecological Economics

journal homepage: www.elsevier.com/locate

Game theory and climate diplomacy

Stephen J. DeCanio [a,*], Anders Fremstad [b]

[a] University of California, Santa Barbara, USA
[b] University of Massachusetts, Amherst, USA

John A. Paravantis, April 2021

There is a long tradition of applying **Game Theory** to problems of **International Relations** (IR)

GEOPOLITICS AND GEOSTRATEGIES: TRENDS AND PERSPECTIVES

POLITICAL DIMENSION OF THE ART OF WAR BY SUN TZU AND ITS GAME-THEORETICAL BACKGROUND

Mateusz HUDZIKOWSKI, PhD *

Game theory incorporates game theoretic models of both the **realist** and the **liberal** views of international politics

- Consider the 4 atomic games

The payoffs of a game can be either **ordinal** (i.e. only a rank ordering is possible) or **cardinal** (i.e. different outcomes can be compared on an absolute scale)

- Cardinal evaluations of outcomes require much stronger assumptions about the utilities of the agents

The essence of many international relations situations can be captured by the simple 2×2 framework

- Great Power rivalry (e.g. US vs China)
- Rich OECD countries (Global North) vs poor developing nations (Global South)

Game-theoretic models provide an elegant formalization of the strategic interactions that underlie international climate negotiations

A number of important **2×2 games** are examined in the paper

- The payoff structure depends on the interpretation of the scientific evidence on climate change

There are (4!)×(4!) = (1×2×3×4)×(1×2×3×4) = 24×24 = 576 ways to arrange four pairs of payoff rankings in an table

- It has been shown that only 576/4 = 144 of these games are distinct
- These 144 distinct games of the simple 2×2 type, may be organized in a unified *"topological framework"*
  - Such a *"New Periodic Table"* of the 2×2 games leads to novel insights about the nature of the games
  - Of these, the paper examines 25 2×2 games that might be relevant to Climate negotiations

The paper organizes subsets of the climate-relevant games into categories (with specific characteristics)

Each player chooses one of two strategies: "**Abate**" or "**Pollute**"

		Column's Strategy	
		Abate	Pollute
Row's Strategy	Abate	a, u	b, v
	Pollute	c, w	d, x

Fig. 1. Generic 2-player game.

The pollutant is **greenhouse gas (GHS) emissions**

Climate relevant restrictions

- Each player's pollution imposes a negative externality on the other
- The outcome (Abate, Abate) is preferred to the outcome (Pollute, Pollute) by both players

Additional restrictions

- There is no economic or geopolitical advantage to be gained by either party if both pollute
- Also neither party’s pollution benefits the other party

These restrictions reduce the number of Climate-relevant 2×2 games (from 144) to 25

The authors disregard the claim that some countries or regions might benefit from Global Warming

–*“Dangerous anthropogenic interference with the Climate”*

+Local increases in agricultural productivity stemming from

Warming

Changes in precipitation patterns

CO_2 fertilization.

The authors also examine a class of *“Cooperate-Defect”* games

- *“Characterized by each player having a dominant preference for a particular strategy by the other player”*
- There are 36 such 2×2 games
- These games are applicable to collective action problems such as arms race games

Important findings of the paper:

- There is no reason to assume from the outset that the Climate problem is inherently one of international **conflict**.
- There are games satisfying the Climate relevant restrictions in which reaching an international **agreement** is relatively easy

Such no-conflict games have the highest payoffs (4,4) for the (Abate, Abate) strategies

Harmony: 366		
	Abate	Pollute
Abate	4, 4 *,+	3, 2
Pollute	2, 3	1, 1
Row's DS = Abate Column's DS = Abate		

No Conflict: 311		
	Abate	Pollute
Abate	4, 4 *,+	2, 3
Pollute	3, 2	1, 1
Row's DS = Abate Column's DS = Abate		

Pure Common Interest: 316		
	Abate	Pollute
Abate	4, 4 *,+	2, 2
Pollute	3, 3	1, 1
Row's DS = Abate Column's DS = Abate		

Pure Common Interest: 361		
	Abate	Pollute
Abate	4, 4 *,+	3, 3
Pollute	2, 2	1, 1
Row's DS = Abate Column's DS = Abate		

326		
	Abate	Pollute
Abate	4, 4 *	1, 2
Pollute	3, 3 +	2, 1
Column's DS = Abate		

362		
	Abate	Pollute
Abate	4, 4 *	3, 3 +
Pollute	2, 1	1, 2
Row's DS = Abate		

321		
	Abate	Pollute
Abate	4, 4 *	1, 3
Pollute	3, 2 +	2, 1
Column's DS = Abate		

312		
	Abate	Pollute
Abate	4, 4 *	2, 3 +
Pollute	3, 1	1, 2
Row's DS = Abate		

* Nash Equilibrium
\+ Maxi-min Equilibrium
DS = Dominant Strategy

Fig. 3. The eight No-conflict Climate-relevant games.

Rational players will settle on the (Abate, Abate) strategy outcome whether they follow Nash or Maxi-min strategies

- The Nash equilibrium and the Maxi-min equilibrium do not have to coincide

If both players play Abate in these games, neither has any incentive to *"defect"* and begin polluting

- Abate is the dominant strategy for both players (in *some* of these games)

Maxi-min strategies are preferred by **risk-averse** players

- This could be a route by which climate stability is reached

There might be a need for international cooperation to make sure the parties understand that (Abate, Abate) is superior to any other outcome, even in the case of no-conflict games

- Turning a **battle-of-the-sexes** type of strategic situation into a **cooperation** game

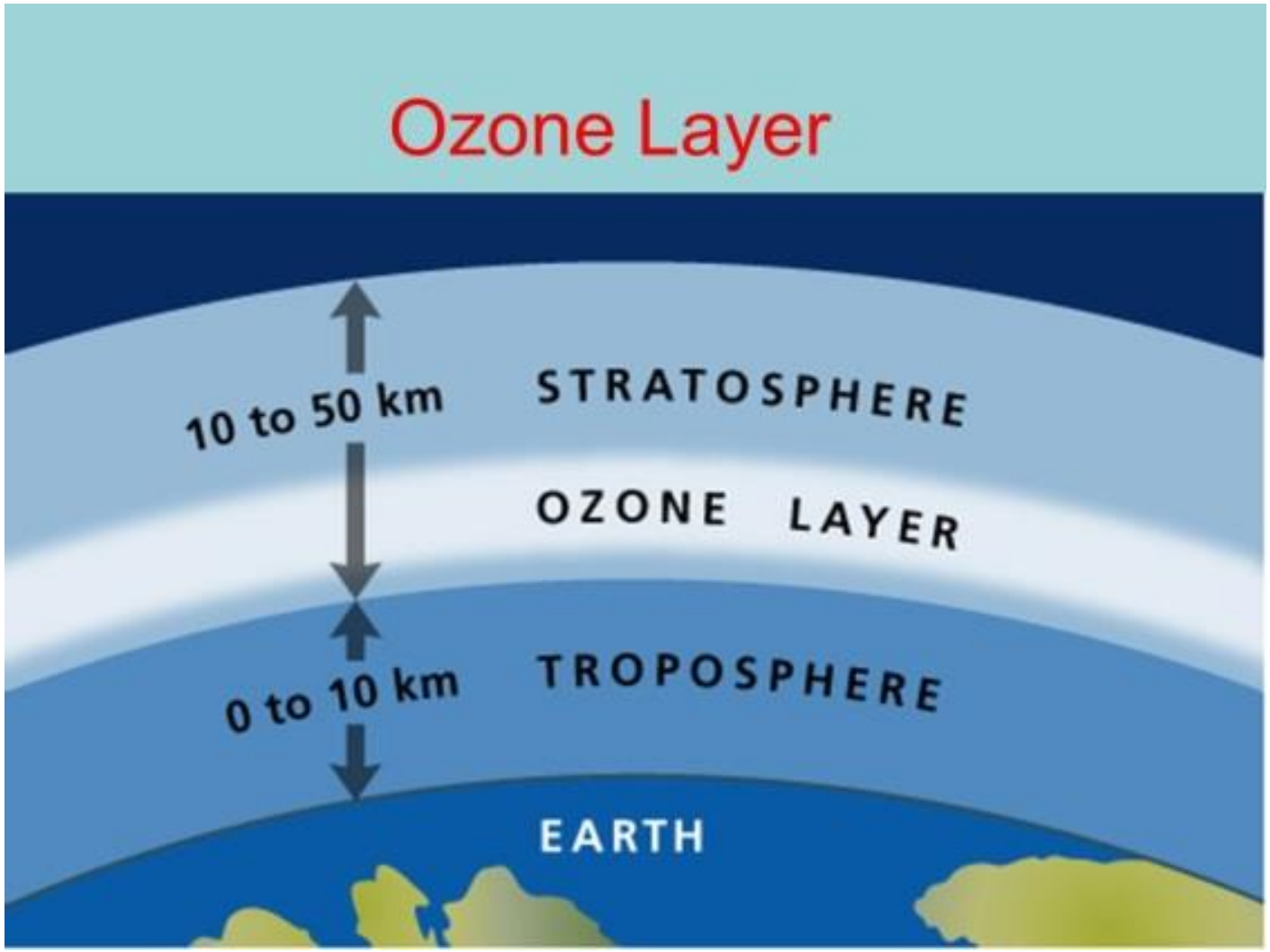
A game-theoretic outcome is **Pareto efficient** (=Pareto optimal) if there is no other outcome that makes every player at least as well off

- A Pareto optimal outcome cannot be improved upon without hurting at least one player
- A Nash equilibrium is considered payoff dominant if it is Pareto superior to all other Nash equilibria in the game
- Often, a Nash Equilibrium is not Pareto efficient implying that the players' payoffs can all be increased

Negotiations could serve to build trust among the parties whose tendency might otherwise be to *"go it alone"*

- When the preferences of the negotiating partners correspond to any of those 8 *"no-conflict"* games, the prospects for an international agreement are good

The success of the *Montreal Protocol on Substances that Deplete the Ozone Layer* (1987) is attributed to its belonging to this category



The damage from stratospheric ozone depletion may have loomed as serious to major countries like the United States

- Abatement of the ozone depleting substances was regarded as a dominant strategy
- Accession to the Montreal Protocol became a dominant strategy for most industrialized countries.

However, the ease of arriving at the Pareto-superior Nash equilibrium in the case of ozone depletion is evidence that this situation does not capture the payoff structure of international climate diplomacy

- Despite three decades of negotiations on climate change, beginning with the run-up to the Framework Convention on Climate Change in Rio de Janeiro in 1992, very little progress has been made.


UNITED NATIONS CONFERENCE ON
ENVIRONMENT AND DEVELOPMENT
Rio de Janeiro 3 – 14 June 1992

The lack of agreement on Climate is like Sherlock Holmes' "*The Adventure of Silver Blaze*", in which the dog famously did not bark (https://en.wikipedia.org/wiki/The_Adventure_of_Silver_Blaze)

Gregory (Scotland Yard detective): “Is there any other point to which you would wish to draw my attention?”

Holmes: “To the curious incident of the dog in the night-time.”

Gregory: “The dog did nothing in the night-time.”

Holmes: “That was the curious incident.”

“I had grasped the significance of the silence of the dog … Obviously the midnight visitor was someone whom the dog knew well.”

— Sherlock Holmes

"…instead of the dog that didn't bark, as expected, my brief described a plaintiff who didn't see a physician for his anxiety and depression and didn't seek Social Security disability income and I asked the Court to draw a conclusion from those negative facts…"

— from an Alabama lawyer on the absence of expected facts (https://brieflywriting.com/2012/07/25/the-dog-that-didnt-bark-what-we-can-learn-from-sir-arthur-conan-doyle-about-using-the-absence-of-expected-facts)

Prisoner's Dilemma (PD) is the best known game of Game Theory.

PD is an example of non-zero sum conflict.

Zero Sum

(http://thejosevilson.com/resisttfa-zero-sum-game)

PD and similar games are only a small subset of the 144 distinct 2×2 games.

In an IR context, a 2×2 PD may be understood by considering the following

- Both countries would benefit from jointly reducing emissions by playing Abate (*"cooperate"*)
- Because abatement of emissions is costly, the worst outcome for either country (at least in the short run, i.e. if the game if not repeated) is to play Abate while the other continues business as usual (i.e. polluting)!

So, PD is characterized by a strong incentive to *"defect"*!

SOME OF MACHIAVELLI'S MAXIMS

- A prince never lacks legitimate reasons to break his promise.
- A wise ruler ought never to keep faith when by doing so it would be against his interests.
- Before all else, be armed.
- Benefits should be conferred gradually; and in that way they will taste better.
- Hatred is gained as much by good works as by evil.

PD is characterized by the existence of a **dominant strategy equilibrium** that is not the same as the **cooperative solution**.

		Column's Strategy	
		Abate	Pollute
Row's Strategy	Abate	3, 3	1, 4
	Pollute	4, 1	2, 2 *,+

* Nash Equilibrium
\+ Maxi-min Equilibrium

Fig. 4. Climate policy as the Prisoner's Dilemma: 111.

Both countries would be better off if they could somehow negotiate an enforceable international **agreement** to Abate.

- This turns out to be elusive because for both countries the highest priority is to prevail in geopolitical competition
- Each would always have an incentive to defect (unilaterally)

In contrast to the outcome of (Abate, Abate), the Nash equilibrium (Pollute, Pollute) is self-enforcing

- Neither country has an incentive to deviate from it (unilaterally).

Both players would prefer a payoff of (3,3) to the Nash equilibrium of (2,2)

- They cannot get there, because they are rational players

How can PD be solved?

- In the global arena, **treaties** could be designed that impose trade sanctions on countries that do not join and comply.
- In the long run, PD may be solved if it becomes **repeated**.
  - If the game were played repeatedly, then a *"tit for tat"* punishment/reward strategy could be employed to train all the parties in the benefits of cooperation

Another way out of the PD is to arrange for **side payments** (or enforceable agreements) to discourage defection.

- These side payments could in fact be drawn from the surplus generated by the difference between the (3,3) and (2,2) outcomes
- They would have to be sufficient to deter defection, and for the cooperative solution to become Nash equilibrium
  - In such a case it would no more be a PD game

If it were known that side payments were possible, each party would have an incentive to announce its intention to defect

- Defectors could be denied green technology transfers.
- Help in one sphere (for example, fighting **terrorism** or suppressing **illegal drug trade**) could be offered in return for cooperation on climate issues

The pressure to comply could take softer forms as well:

- Polluters could be subject to opprobrium and shame if they did not conform to the preferences of the world community
  - (Unrealistic)

IR perspective

- Great Powers are concerned primarily with their **relative** (rather than **absolute**) power
  - A country might reject any treaty that provided greater **gains** for its rival than for itself.

Game theoretic models would have to be set up, so that their ordinal payoffs accounted for this.

To recap PD, the general idea is

- Bringing in **rewards** and **punishments** that are outside the structure of the game's payoffs

The PD is related to the concept of **public goods** (such as pollution abatement)

- They are provided within States by enforcing environmental laws with civil and criminal penalties for those who pollute

Table 3.1 Non-rivalry and non-excludability

	Rival	Non-rival
	Pure private goods	**Mixed goods**
Excludable	• Cornflakes • Cars • Chocolate	• Pay per view TV • Toll bridges • Private roads
	Mixed goods	**Pure public goods**
Non-excludable	• State education • Public health • Open access resources such as ocean fishing fields, city streets and town parks	• National defence • Lighthouses • A clean environment • Very large national parks

Carmichael, F. (2005): *A Guide to Game Theory*. Prentice Hall Financial Times

The Tragedy of the Commons

The population problem has no technical solution; it requires a fundamental extension in morality.

Garrett Hardin

At the end of a thoughtful article on the future of nuclear war, Wiesner and York (*1*) concluded that: "Both sides in the arms race are . . . confronted by the dilemma of steadily increasing military power and steadily decreasing national security. *It is our considered professional judgment that this dilemma has no technical solution.* If the great powsional judgment. . . ." Whether they were right or not is not the concern of the present article. Rather, the concern here is with the important concept of a class of human problems which can be called "no technical solution problems," and, more specifically, with the identification and discussion of one of these.

It is easy to show that the class is not

What Shall We Maximize?

Population, as Malthus said, naturally tends to grow "geometrically," or, as we would now say, exponentially. In a finite world this means that the per capita share of the world's goods must steadily decrease. Is ours a finite world?

A fair defense can be put forward for the view that the world is infinite; or that we do not know that it is not. But, in terms of the practical problems that we must face in the next few generations with the foreseeable technology, it is clear that we will greatly increase human misery if we do not, during the immediate future, assume that the world available to the terrestrial human population is finite. "Space" is no escape (*2*).

A finite world can support only a finite population; therefore, population growth must eventually equal zero. (The case of perpetual wide fluctuations above and below zero is a trivial variant that need not be discussed.) When this condition is met, what will be the situa-

Hardin, G. (1968). *The Tragedy of the Commons*. Science, New Series, Vol. 162, No. 3859, pp. 1243-1248

There are other games in the PD “neighborhood”, i.e. with payoff structures similar to the PD and its logic

Alibi: 412		
	Abate	Pollute
Abate	3, 4	1, 3
Pollute	4, 1	2, 2 * +
Row’s DS = Pollute		

Alibi: 221		
	Abate	Pollute
Abate	4, 3	1, 4
Pollute	3, 1	2, 2 * +
Column’s DS = Pollute		

* Nash Equilibrium
+ Maxi-min Equilibrium
DS = Dominant Strategy

Fig. 5. Climate games similar to the Prisoner's Dilemma.

Fossil fuels still constitute the bulk of the world's energy supply, and **energy** is central to modern industrial power.

- Interventions in the energy sector strikes close to the heart of an industrialized nation's economic strength.

From an IR perspective, countries with global influence fear that they would be weakened if they were required to

- Scale back their energy production and consumption
- Substitute more expensive primary energy sources for fossil fuels

The **Coordination Game** (CG) is another environmentally relevant (rather than just climate-relevant) game that

- Shows the same kind of priority given to geopolitical competition as the PD games
- Offers a greater possibility for international cooperation

		Column's Strategy	
		Abate	Pollute
Row's Strategy	Abate	4, 4 *	1, 3
	Pollute	3, 1	2, 2 *,+

* ***Two*** Nash equilibria

\+ Maxi-min equilibrium

Fig. 6. Climate policy as a Coordination problem: 322.

In a game theoretic setting

- Both countries are highly averse to playing Abate while their rival plays Pollute
- The best outcome for both is to play Abate jointly
- The incentive to defect is eliminated

Still, in an IR context, an **agreement** by both parties to play Abate must first be reached

The (Pollute, Pollute) outcome, which is Pareto-inferior, entails a risk that the planet is damaged.

- At least *"We Will All Go Together When We Go"* (as in the Tom Lehrer 1959 song that refers to nuclear war and *"complete participation in that grand incineration"*).
- https://youtu.be/TIoBrob3bjI

In the CD there are **two Nash equilibria** (both countries either play Abate or Pollute)

- One country playing Abate while the other plays Pollute leads to losing out in the short-run geopolitical competition
- The **diplomatic** problem here is one of **equilibrium choice**

In the rational world of IR, the inferior equilibrium (Pollute, Pollute) should be readily disregarded in favor of the payoff dominant (Abate, Abate) which is Pareto optimal

- So, an international agreement to Abate should be self-reinforcing

Both the PD and the CG exhibit the classic **collective action** problem

- The worst outcome for a player is to abate (i.e. contribute to paying for the public good) while the other player is a **free rider**.

The difference is that

- In the **Coordination Game**, the highest-valued outcome for both parties is achieved when they cooperate
- In the **Prisoner's Dilemma** the best outcome for a party is to Pollute while the other Abates

Is global climate protection more like a PD or a CD?

The answer depends on how severe the risk of catastrophic climate change is perceived as being

If countries read the science as saying that climate change is an **existential threat** to humanity and civilization, then the world is in a CG

If not even the risk of extinction is more important than gaining **geopolitical advantage** over the other competing powers, the PD characterizes the situation and the outlook for cooperation is dim

The authors assert that the overriding barrier to achieving an international agreement to protect the climate may be a failure of the leading governments to grasp the seriousness of the climate risk

- Perhaps at a future time, when the science is crystal clear and indisputable, there will be a relatively sudden realization of the magnitude of a global environmental risk.
- This behavior appears to be consistent with the history of the negotiation and subsequent success of the Montreal Protocol

Wrapping up, here are some more games in the CG neighborhood:

Coordination games that are not climate-relevant

Pure coordination: 334

	Abate	Pollute
Abate	4, 4*	1, 1
Pollute	2, 2+	3, 3*

Pure coordination: 343

	Abate	Pollute
Abate	4, 4*	2, 2+
Pollute	1, 1	3, 3*

Coordination: 333

	Abate	Pollute
Abate	4, 4*	1, 2
Pollute	2, 1	3, 3*,+

Coordination: 344

	Abate	Pollute
Abate	4, 4*,+	2, 1
Pollute	1, 2	3, 3*

Asymmetric coordination: 323

	Abate	Pollute
Abate	4, 4*	1, 2
Pollute	3, 1	2, 3*,+

Asymmetric coordination: 332

	Abate	Pollute
Abate	4, 4*	1, 3
Pollute	2, 1	3, 2*,+

Asymmetric coordination: 324

	Abate	Pollute
Abate	4, 4*	1, 1
Pollute	3, 2+	2, 3*

Asymmetric coordination: 342

	Abate	Pollute
Abate	4, 4*	2, 3+
Pollute	1, 1	3, 2*

*Nash equilibria.
+Maxi-min equilibrium.

The **Chicken Game** is also relevant to climate change.

		Column's Strategy	
		Abate	Pollute
Row's Strategy	Abate	3, 3 +	2, 4 *
	Pollute	4, 2 *	1, 1

* Nash Equilibria
+ Maxi-min equilibrium

Fig. 7. Climate negotiations as the game of "Chicken": 122.

Chicken has two Nash equilibria – one party Pollutes while the other Abates

- Agents may operationalize their risk aversion by selecting Maxi-min strategies
- **Risk-averse** parties might reach the (Abate, Abate) outcome rather than the Nash equilibria

There are four other climate-relevant games having the same logic as Chicken

426		
	Abate	Pollute
Abate	3, 4 +	2, 2
Pollute	4, 3 *	1, 1
Column's DS = Abate		

262		
	Abate	Pollute
Abate	4, 3 +	3, 4 *
Pollute	2, 2	1, 1
Row's DS = Abate		

421		
	Abate	Pollute
Abate	3, 4 +	2,3
Pollute	4, 2 *	1, 1
Column's DS = Abate		

212		
	Abate	Pollute
Abate	4, 3 +	2, 4 *
Pollute	3, 2	1, 1
Row's DS = Abate		

* Nash equilibrium
+ Maxi-min equilibrium
DS = Dominant Strategy

Fig. 8. Asymmetric games like Chicken.

As in the case of the CD, one may imagine negotiations as serving the function of persuading the country inclined to pollute of the **severity of risks** posed by uncontrolled climate change

- In other words, the issue in Chicken is to convince all countries that it is in their interest to be risk-averse players
  - Only solid scientific evidence may achieve this, and even then probably at the last minute

The problem remains that the defecting party can exploit the risk averseness of the Maxi-min player

- Unfortunately, while playing a Maxi-min strategy can be attractive to a risk-averse player, it may allow a more ruthless competitor to gain the advantage by **defecting** from the Maxi-min equilibrium

In the (“*profoundly*”) **unhappy games**, both the Nash equilibrium and the Maxi-min agreement are for one country to pollute while the other abates

121		
	Abate	Pollute
Abate	3, 3	2, 4 *,+
Pollute	4, 1	1, 2

112		
	Abate	Pollute
Abate	3, 3	1, 4
Pollute	4, 2 *,+	2, 1

The Four Climate-Relevant “Type” Games

261		
	Abate	Pollute
Abate	4, 3	3, 4 *,+
Pollute	2, 1	1, 2

211		
	Abate	Pollute
Abate	4, 3	2, 4 *,+
Pollute	3, 1	1, 2

416		
	Abate	Pollute
Abate	3, 4	1, 2
Pollute	4, 3 *,+	2, 1

411		
	Abate	Pollute
Abate	3, 4	1, 3
Pollute	4, 2 *,+	2, 1

* Nash equilibrium

\+ Maxi-min equilibrium

Fig. 9. The six “unhappy” Climate-relevant games.

In these unhappy games, players

- have asymmetric preferences
- inhabit different moral universes
  - Think of the **US** and the **European Union**

Unhappy games are appropriate for

- Exploring morally ambiguous situations
- Players debating morality from fundamentally different material situations

Visualize the climate debate in these terms

- Listening to the **rich developed** nations and the **poor developing** countries talk past each other

In such a case

- The prospect of reaching an agreement is bleak
- Climate stabilization could be achieved only by some combination of carrot and stick
  - (Side payments and coercion)

Finally, let's examine games with no pure-strategy Nash equilibrium (also called **Cycle games**).

The lack of a Nash equilibrium makes these games somewhat futile

Cycle: 422		
	Abate	Pollute
Abate	3, 4	2, 3 +
Pollute	4, 1	1, 2

Cycle: 222		
	Abate	Pollute
Abate	4, 3	1, 4
Pollute	3, 2 +	2, 1

+ Maxi-min equilibrium

Fig. 10. Games with no Pure-Strategy Nash Equilibrium.

The cyclical nature implies that

If both players started with the (Pollute, Pollute) outcome

Row would switch to Abate

Column would switch to Abate

Row would switch to Pollute

and Column would switch to Pollute, so it would be back at the (Pollute, Pollute) outcome!

An international agreement with incentives and/or enforcement provisions outside the payoff matrix would be necessary to maintain (Abate, Abate).