Ecological Economics 85 (2013) 177–187

Contents lists available at ScienceDirect

Ecological Economics

journal homepage: www.elsevier.com/locate/ecolecon

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

ARTICLE INFO

Article history:
Received 2 December 2010
Received in revised form 18 April 2011
Accepted 25 April 2011
Available online 22 June 2011

Keywords:
Climate change
Game theory
Nash equilibrium
Maxi-min equilibrium
Negotiations
Strategy

ABSTRACT

Starting with the "New Periodic Table" (NPT) of 2×2 order games introduced by Robinson and Goforth (2005), we provide an exhaustive treatment of the possible game-theoretic characterizations of climate negotiations between two players (e.g., Great Powers or coalitions of states). Of the 144 distinct 2×2 games in which the players have strict ordinally ranked utilities, 25 are potentially relevant to climate problem. The negotiations may be characterized as a No-Conflict Game, Prisoner's Dilemma, Coordination Game, Chicken, Type Game, or Cycle, depending on the payoff matrix. Which game corresponds to the actual state of the world depends both on the severity of risks associated with climate change and the perceptions of the governments engaged in the negotiations. Nash equilibrium or Maxi-min equilibrium (or neither) may be the outcome. Achieving universal abatement of greenhouse gas emissions may require side payments or enforcement mechanisms outside the game framework, but we show how the negotiations themselves may offer opportunities to select between Nash equilibria or alter the payoff rankings and strategic choices of the players. In particular, scientific information pointing to the severity of the risks of climate change suggests characterization of the negotiations as a Coordination Game rather than a Prisoner's Dilemma.

1. Background

Game-theoretic models provide an elegant formalization of the strategic interactions that underlie the international climate negotiations. Needless to say, there is a long and lively tradition of applying game theory to problems of international relations, including global environmental protection. We will not attempt to give a comprehensive survey of this literature.[1] Instead, we will comprehensively examine all of the 2×2 order games[2] that might be relevant to the climate negotiations, and that show how the payoff structure depends on interpretation of the scientific evidence. We will argue that assessment of the magnitude of the global climate risk is the key determinant of the kind of "game"[3] being played. This in turn affects the feasibility of reaching an agreement, and the possible role of equity considerations in facilitating an agreement.

Game theory incorporates key elements of both the realist and liberal views of international politics (Stein, 1990). It is consistent with realism because the players are assumed to have a unitary will, that is, each government acts as a single agent rather than as some kind of complex organization whose decisions result from domestic political interactions.[4] At the same time, it shows how self-interested behavior can lead to order and welfare-improving outcomes (though it need not necessarily do so), just as the market economy can. The game-theoretic approach does require that governments are able to

* Corresponding author at: 29 Sassafras Lane, Amissville, VA 20106, USA.
E-mail addresses: decanio@econ.ucsb.edu (S.J. DeCanio), anders.fremstad@gmail.com (A. Fremstad).

[1] A wide-ranging application of game theory to global environmental protection (which also gives a wealth of historical and institutional background for several examples, including the climate negotiations) is Barrett (2003). Stein (1990) offers numerous applications to a range of situations in international relations, and is a fine example of the richness of the 2×2 game framework as a source of insight into the strategic interactions that can arise.

[2] Order games are games in which the outcomes stemming from governments' policy choices are ranked ordinally (without indifference).

[3] The terminology "game theory" is an historical accident. As Shubik (1983, p.7) put it, "[p]erhaps the word 'game' was an unfortunate choice for a technical term. Although many rich and interesting analogies can be made to Bridge, Poker, and other parlor games, the usual sense of the word has connotations of fun and amusement, and of removal from the mainstream and the major problems of life. These connotations should not be allowed to obscure the more serious role of game theory in providing a mathematical basis for the study of human interaction, from the viewpoint of the strategic potentialities of individuals and groups." There is no way to go back in history and persuade von Neumann and Morgenstern, the intellectual giants whose book (1944) and prior work (von Neumann, 1928) launched the field, to adopt more descriptive titles.

[4] Realist thinking in international relations encompasses more than this principle alone. For example, some Realist theorists emphasize the importance of *relative* power as a priority of governments. For a critical review of Realism see Donnelly (2000).

0921-8009/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolecon.2011.04.016

rank-order outcomes in a manner that is consistent with agent rationality (i.e., a ranking can be assigned to each outcome and the rankings are transitive). Note that the perceived interests of the governments can allow for some weight being given to the well-being of other nations; all that is required is that the outcomes be ranked. In general, the payoffs of a game can be either ordinal (only a rank ordering is possible) or cardinal (different outcomes can be compared on an absolute scale, such as in monetary units). We will focus most of our attention on ordinal rankings. Conclusions based only on ordinal rankings are more general; cardinal evaluations of outcomes require much stronger assumptions about the utilities of the agents. Ordinal ranking of outcomes allows us to bypass the comparisons of utility across countries with very different levels of income that plague conventional cost–benefit analysis.[5] Because ordinal rankings are preserved under any positive monotonic transformation, any conclusions based only on the ordinal ranking of outcomes carry over if the game payoffs are expressed in monetary or utility units.

Consider first the simplest possible games. There are only two players, who will be identified as "Row" and "Column." We will consider games with more players below, but we concur with Barrett that the essence of many international relations situations can be captured by the simple 2×2 framework.[6] We use this nondescript "Row" and "Column" terminology for the players because the games will be interpreted as representing different kinds of international relations — sometimes Great Power rivalry (as between the United States and China, for example) and sometimes other strategic interactions (as between the relatively rich OECD countries and the relatively poor developing nations). Each player chooses one of two strategies, "Abate" or "Pollute."

The payoffs of the games are given in a simple matrix, each cell of which has two elements, the payoff for Row followed by the payoff for Column. The generic payoff matrix is shown in Fig. 1. Thus, a in the upper left cell of the matrix is the payoff for Row if Row chooses the strategy Abate, and Column chooses the strategy Abate. Similarly, u is Column's payoff from this pair of strategy choices. The payoffs to each player are measured in ordinal terms, so $\{a, b, c, d\}$ and $\{u, v, w, x\}$ can take on values {4, 3, 2, 1}, with 4 corresponding to the best outcome, 3 and second-best outcome, 2 the next-to-worst outcome, and 1 the worst outcome for each player.

There are 144 distinct games of this simplest type. The number of distinct 2×2 order games has been known since the 1960s, but recent work by Robinson and Goforth (2005) shows that these games can be organized in a unified topological framework based on a natural measure of the "distance" between payoff structures.[7] Robinson and Goforth's book synthesizes what has been known about the classification of 2×2 games to date, and their "New Periodic Table" (NPT) of the 2×2 games efficiently organizes the information and leads to new insights about the nature of the games. We will use the

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

Fig. 1. Generic 2-player game.

Robinson and Goforth NPT to provide an *exhaustive and theoretically unified treatment of all 2×2 games that might be relevant to the climate negotiations*. We will be able to cover every possible case of preferences and strategic interactions, and drawing on the NPT topology we will show how different subsets of the climate-relevant games fall into categories with specific characteristics. Determining just which situation is most descriptive of the actual state of play in the negotiations then depends on how the preference rank-orderings of the players are assessed.

First we must establish which of the 2×2 games are potentially applicable to the climate problem. We narrow down the number of games by requiring that the payoff structures satisfy two "climate relevant" restrictions: (1) The outcome (Abate, Abate) is preferred by both players to the outcome (Pollute, Pollute), and (2) Each player's pollution imposes a negative externality on the other. The first of these two restrictions amounts to assuming that there is no economic or geopolitical advantage to be gained by either party if both pollute instead of both abating, and that the climate problem is real. It does not require that climate is either party's top priority. The second restriction amounts to the presumption that neither party's pollution benefits the other party. In the generic payoff matrix of Fig. 1, the first restriction says that $a>d$ and $u>x$. The second restriction requires $a>b$, $c>d$, $u>w$, and $v>x$. These two restrictions reduce the number of climate-relevant 2×2 games to 25.[8]

It should be noted that our two climate-relevance conditions apply to countries' greenhouse gas emissions. A small country rich in oil or gas reserves may derive much of its national income from the export of its fossil fuel resources. From a short-term perspective, the government of such a country might prefer that the rest of the world adopt the Pollute strategy, violating our "negative externality" condition (2). This situation might apply to a few countries, but not to the major powers.[9]

In addition, we do not consider the thinly-supported claims that some countries or regions would benefit from global warming (see the critique of these claims in Ackerman et al. (2009a)). The world is already committed to some amount of warming because of cumulative emissions to date, and the pending policy question is how much more warming can be allowed if we are to avoid "dangerous anthropogenic interference with the climate," even if there may be some relatively minor increases in agricultural productivity in a few regions stemming from the warming, changes in precipitation patterns, and CO_2 fertilization.

[5] For example, global welfare-maximizing integrated assessment models have to employ a scheme such as Negishi weighting to construct the social welfare function to be maximized. If Negishi weights are used, the social welfare function embodies and largely freezes the current distribution of income (Stanton, 2010).

[6] Barrett (2003) cites Stein that "[m]ost basically, nations choose between cooperation and conflict, and such situations underlie the entire range of international relations from alliances to war" (1990, pp. 3–4).

[7] There are (4!)×(4!) = 576 ways to arrange four pairs of utility rankings in an array such as Fig. 1. Robinson and Goforth show that only 576/4 = 144 of these games are distinct (2005, Chapter 2, pp. 15–19). Rapoport and Guyer (1966) and Brams (1977) count only 78 distinct games because they define "'distinct' in the sense that no interchange of the column strategies, row strategies, players, or any combination of these can turn one game into another — that is, these games are structurally different with respect to these transformations" (Brams, 1983, p. 173), and they eliminate "reflections" as defined by Rapoport and Guyer. However, "[t]here are strong arguments against eliminating reflections" (Robinson and Goforth, 2005, p. 19) if the players are not indistinguishable. Thus, for our purpose (determining the games that might be applicable to climate negotiations) the Robinson and Goforth count of 144 is appropriate.

[8] Of the games having essentially different structures, some are analyzed so frequently as to be given distinctive names, such as the Prisoner's Dilemma, Chicken, Battle of the Sexes, and Stag Hunt. We will use the names given by Robinson and Goforth if they seem apposite. Some of the games are named after Cold War or Vietnam-era events or situations that have nothing to do with climate, and in these cases we identify the games by their NPT numbers only.

[9] An international climate agreement that reduced the risk of climate change would correct a market failure (allowing free disposal of greenhouse gases into the atmosphere) but would create a pecuniary externality by causing a loss of fossil fuel wealth and revenues to OPEC members. Pecuniary externalities are typically considered to be part of the dynamic market process and are in fact necessary for allocational efficiency. The political process, however, makes no distinction between pecuniary externalities and "real" externalities (Holcombe and Sobel, 2001).