

SYMPORIUM
A CHANGING MORAL CLIMATE



GAME THEORY AND THE ETHICS OF
GLOBAL CLIMATE CHANGE

BY RONY SMEAD AND RONALD SANDLER

[THIS PAGE INTENTIONALLY LEFT BLANK]

Game Theory and The Ethics of Global Climate Change

Rory Smead

&

Ronald Sandler

Stephen Gardiner begins *A Perfect Moral Storm* with the claim that “Sometimes the best way to make progress in solving a problem is to clarify what the problem is” (3).¹ One of his goals in the book is to do just this for global climate change—or, more precisely, for the problem of why it has been so difficult to generate national policies and international agreements to address it. He believes that global climate change is the “perfect moral storm” in that it has several features that make generating action to address it extremely difficult—for example, it is intergenerational, global in scale, involves difficult to quantify values, and is multi-level (e.g. international and national). His purpose is “to get clearer about” these features of the problem “as a preliminary to generating and assessing potential solutions” (4). Central to this project is game theory. Gardiner employs several different game theoretic approaches—e.g. Prisoner’s

¹ Stephen Gardiner, *A Perfect Moral Storm. The Ethical Tragedy of Climate Change* (Oxford: Oxford University Press 2012). Unless otherwise specified, parenthetical references refer to this text.

Dilemma and Battle of the Sexes—to try to shed light on the problem and to explain why it is so difficult to address.

In this commentary, we focus on Gardiner's use of game theory to illuminate the problem of generating responsiveness to global climate change. Gardiner is quite up front that he does not provide a “full account of the role of game theory in ethical analysis” (50). So this commentary addresses a part of his project that is not fully articulated within *A Perfect Moral Storm*, but that is crucial to it. We agree with Gardiner that clarifying a social problem is often a crucial prelude to effectively addressing it. We also agree with him that game theory can play a valuable role in ethical analysis, and that one of its functions is helping to characterize problems. We further agree with him that the issue of why there has been so little responsiveness to global climate change is one that game theory can help to illuminate. However, we think that game theory can play other, more robust roles as well. We also have some concerns about Gardiner's general approach to using game theory to illuminate the problem of global climate change, as well as some concerns about his uses of particular games. In Section 2, we focus on the relationship between game theory and ethics, as well as on Gardiner's general approach to applying game theoretic analyses to climate change. In Section 3, we discuss some particular games that Gardiner uses and how he uses them, as well as some games that he might have employed but did not. Overall, we find that much of Gardiner's use of game theory is both appropriate and helpful. However, in some other respects Gardiner is under-utilizing game theoretical analysis or not using the relevant games in the most effective ways.

II

Ethics and Game Theory

In *A Perfect Moral Storm*, Gardiner's approach to using game theory to characterize the problem of climate change is to try to find the game that is descriptively accurate. By that we mean he compares the structure of various games—particularly with respect to player positions and payoff possibilities—to the situation of climate change “players” (e.g. nations, businesses, and individuals) to see if the game accurately represent the player’s positions relative to each other and the problem. He does this, at some length, with the Prisoner’s Dilemma (55ff; 104ff), Battle of the Sexes (87ff), and tragedy of the commons (108ff), for example. The idea is that if we can match the right game theoretic structure to the problem of climate change, then we can apply what we know of that game to better understand why it is that so little progress has been made in addressing climate change. Moreover, we can begin to explore methods for finding positive solutions or outcomes within those games to see if they might apply to the problem of climate change. This is a fruitful approach to using game theory in ethics. It allows us to: 1. Identify key features of the problem; 2. Characterize possible solutions spaces; 3. Explore possible strategies for getting us to desired solutions.

In these respects, game theory is not prescriptive. It does not provide a method of determining what solutions are desirable (or ethically preferable) or what means are acceptable for accomplishing them. Nor does it tell us which features of the problem are ethically salient; or whether (or how urgently) we ought to work to solve the problem in the first place. What it can do is describe which features of the problem make it difficult to solve; help to determine whether solutions are possible; and help to identify methods or mechanisms (e.g. incentive changes) that

can facilitate movement toward certain solutions. In *A Perfect Moral Storm*, Gardiner is focused on the first of these: the features of the problem that make it difficult to solve. The question, after all, is why, if global warming is such a large problem and we have been working on it for two decades, so little has been accomplished (116). So while Gardiner's use of game theory to characterize the problem is appropriate, it is not all that game theory can contribute to descriptive ethics. Moreover, it might not even be the most valuable role it can play. After all, in this role it is not so much that the games illuminate the problems, but rather the problems are matched to the appropriate games, which if done well, requires prior knowledge of both the problems and the games. The value, or how the games bring in new information on the problem, arises when we look at how the games that are operative can help illuminate possible solutions and pathways to them. Gardiner appears sympathetic with this—after all, he says that characterizing the problem is prelude to exploring solutions.

However, it is important in evaluating which games are descriptively accurate for climate change that we be clear what game theory does and does not (and can and cannot) characterize. First, no game is going to be a comprehensive characterization of the problem, since the games are at best idealizations for only particular features of an actual problem. This is particularly so for a problem like climate change, where there are multiple levels (e.g. international, national, and personal) and multiple types of “players” or agents (e.g. governments, corporations, and individuals). Moreover, even if we could represent the problem in one big game, this misses part of the point of game theory—its ability to simplify strategic interactions so that their underlying principles can be understood. We don’t want a fully realistic game because that would be just as hard as the real climate situation to analyze. Given this, it should not be taken as problematic if a game “fails as a *general* account of the problem posed by global

climate change” (124). The descriptive use of climate change should not involve asking ‘Which game theoretic is the right one for climate change?’ as Gardiner often seems to (e.g. 106), but rather ‘Which games can usefully characterize which aspects of the climate problem?’

Furthermore, it should not be taken as problematic if there are aspects of the climate problem that particular games fail to characterize. For example, Gardiner is critical of the tragedy of the commons analysis of climate change because it does not adequately capture certain moral dimensions of the problem. As he puts it, “I conclude that the basic tragedy of the commons model is, at best, seriously incomplete. The model’s neglect of differences in vulnerability, and especially the plight of the global poor, means that it obscures vital features of the problem at hand” (122). But game theoretic analysis (including tragedy of the commons) is not meant to capture those features of the problem—it is a decision theoretic model—and it would only obscure other aspects of the problem, including moral ones, if it were supposed that it did constitute a full characterization of the problem. Moreover, to restate a point made above, game theoretic analyses do not provide assessments of which problems are ethically important, which solutions are ethically preferable, or which approaches to generating solutions are ethically acceptable. Therefore, it should be no surprise that game theory, let alone a particular game, often fails to capture all morally salient features of the climate problem.

That said, there are games that have been used to study fairness, such as the Nash bargaining game or the Ultimatum game.² These games deal with distribution of goods rather than

² K. Binmore, *Natural Justice* (Oxford: Oxford University Press 2005); B. Skyrms, *The Evolution of the Social Contract* (Cambridge: Cambridge University Press, 1996)

achieving some cooperative goal. For reasons already discussed, they do not capture the full complex nature of the climate crisis, or all the morally salient features of it. However, they represent one important aspect of solving the problem—any agreement on mitigation is going to involve questions of distribution. Games of cooperation and social dilemmas have received the majority of attention in game theoretic approaches to climate change, but other types of games should also be considered (we discuss this point again below). Some games might be useful for characterizing certain aspects of international climate negotiations (e.g. bargaining games), while others are useful for characterizing competition over natural resources (e.g. Chicken) or simply coordination to achieve a common good (e.g. the Stag Hunt). Again, rather than attempt to identify the game that is descriptively adequate for representing the climate problem, a more effective approach would seem to be a pluralistic one that uses multiple games to illuminate different aspects and dynamics of a complicated social situation.³

III

Game Theory and Climate Change

Much of the game theoretic analysis in *A Perfect Moral Storm* involves the well-known Prisoner's Dilemma and various extensions of it—e.g. public goods games and asymmetric versions. There is good reason for this. The Prisoner's Dilemma and similar games represent cases of deep and serious conflict between individual self-interest and the collective good. Everyone

³ See S.J. DeCanio and A. Fremstad, “Game theory and climate diplomacy,” *Ecological Economics* 85 (2013), 177-187, for an overview the ways in which all two-player, two-strategy games may be relevant to climate change.

would be better off if everyone cooperated, but it is in no one's self-interest to do so, even if everyone else is going to cooperate. Because cooperation is irrational from the perspective of individual self-interest, it does not form an equilibrium of the game and cooperative states are unstable. Thus, the Prisoner's Dilemma represents the "hard case" for solving a cooperation problem: if you can solve the problem of cooperation in the Prisoner's Dilemma, you can solve it in a number of other games as well. Furthermore, there are several important aspects of climate change, such as mitigating carbon emissions, which seem to mirror a Prisoner's Dilemma.

However, the important decision-theoretic features of the climate problem are not exhausted by the Prisoner's Dilemma. Indeed, Gardiner acknowledges the limitations and incompleteness of this model in many places. Gardiner also employs other games, such as the Battle of the Sexes, which is a game of coordination where each player has a different preference over where to coordinate. The Battle of the Sexes represents a more promising situation than the Prisoner's Dilemma: cooperation is stable and rationally justifiable. But, Gardiner argues that "none of the main claims of the broader Battle of the Sexes model...seems likely to be true of climate change," and consequently that "we should look elsewhere for a compelling account of the shape of the global storm" (102). We agree that Battle of the Sexes does not adequately represent the climate problem (i.e. it is not descriptively accurate), but do not think it should be entirely dismissed because of this. Games such as the Battle of the Sexes should not be viewed as an alternative to the Prisoner's Dilemma (and related games), but rather as a way of representing different aspects of the problem. In this case, this game can illuminate an important obstacle for cooperative endeavors that is missed in the Prisoner's Dilemma: the Battle of the Sexes illustrates what is known as the "equilibrium selection

problem” in game theory.⁴ Most games are not like the Prisoner’s Dilemma, which has a unique (and bad) equilibrium or “solution.” Most games have multiple equilibria and the question becomes, which one should we expect to see? And how do we facilitate convergence to some preferred equilibria over the other possibilities? Consequently, even if we somehow solve the Prisoner’s Dilemma (by changing the game in some way) and make cooperation a possible stable solution, there is still the problem of *reaching* that positive solution.

These points are perhaps best illustrated by another game, the Stag Hunt. In this game two players in a joint stag hunt are simultaneously faced with a decision to stick to their posts (cooperate) or to abandon their post and hunt hare (defect). Hunting hare provides a moderate but certain payoff independent of the actions of others. Hunting stag provides a high payoff if the other player cooperates and nothing if the other player defects. In this game, cooperation is one equilibrium and non-cooperation is another: cooperation is a possible solution to this game (unlike the standard Prisoner’s Dilemma), as is non-cooperation. The cooperative equilibrium is optimal but risky because everyone loses the gains if one player fails to cooperate. The non-cooperative equilibrium is suboptimal but risk free. Gardiner only briefly considers the Stag Hunt (114, footnote 14) and dismisses it for not capturing the climate problem as well as, say, the Prisoner’s Dilemma.

However, if we abandon the idea of trying to find the game that best represents the climate problem, and instead focus on what each game might illuminate about the problem, the Stag Hunt should not be so quickly dismissed. It represents a different,

⁴ C. Harsanyi and R. Selton, *A General Theory of Equilibrium Selection in Games* (Cambridge, MA: MIT Press, 1988); L. Samuelson, *Evolutionary Games and Equilibrium Selection* (Cambridge, MA: The MIT Press 1998).

but also crucial, obstacle to social cooperation. The Prisoner’s Dilemma represents the problem of stability, whereas the Stag Hunt represents a case where there are stable preferable states, but they are hard to reach, since we are “stuck” in a suboptimal but equally rational solution (from the view of individual self-interest). Furthermore, many of the potential “solutions” to the Prisoner’s Dilemma essentially involve changing the incentives and payoff structure of the game, thereby transforming it into a Stag Hunt.⁵ That is, standard “solutions” to the Prisoner’s Dilemma solve the problem of stability of cooperation, but do not solve the equilibrium selection problem; they do not tell us how to get from non-cooperation to cooperation, only that once we achieve cooperation it will persist. The Stag Hunt represents this particular problem in a simple and precise way. Therefore, both the game and the problem it illustrates are valuable to characterizing the problem of generating responsiveness to global climate change. That the game is dismissed and the problem of multiple stable states is not highlighted in *A Perfect Moral Storm*, where the core project is to characterize the problem of climate change using game theoretics, seems to us to be an important oversight.

There is another aspect of game theory that we believe is importantly relevant to the climate issue, but is not discussed in *A Perfect Moral Storm*. Evolutionary game theory is used to represent dynamic settings where (boundedly rational or non-rational) individuals learn and adapt over time.⁶ The motivation for this approach includes the equilibrium selection problem, along with the observation that real agents are less than fully rational. These

⁵ B. Skyrms, *The Stag Hunt and the Evolution of Social Structure* (Cambridge: Cambridge University Press, 2004).

⁶ See B. Skyrms, *The Evolution of the Social Contract* (Cambridge: Cambridge University Press, 1996), for an overview on applications of evolutionary game theory to topics in descriptive ethics.

considerations have led game theorists to be concerned not just with what solutions a game has, but whether and under what conditions the solutions can be reached.⁷ Considering the way in which boundedly rational individuals may adapt or change over time is important if we would like to understand when such agents might be able to reach certain solutions and when they might not. Gardiner's game theoretic analysis is largely focused on the equilibria or "solutions" to the games and not on the way in which behavior within games may change over time. As we discussed earlier, if we are too focused on the stability of cooperation in solving the climate problem, we may miss the other barriers to achieving cooperation and other ways in which we might be able to facilitate reaching positive solutions.

IV

Conclusion

This commentary has focused primarily on the ways in which Gardiner's use of game theory to help to understand and develop responsiveness to global climate change could have been expanded or improved. So we should like to conclude by emphasizing again that we believe that Gardiner has the big methodological picture right—it is crucial to solving this problem that we understand it better and game theoretic analysis is a valuable tool in doing so. We also believe that his discussion of the problem in light of the Prisoner's Silemma and tragedy of the commons, in particular, makes a very large contribution in this regard. *A Perfect Moral Storm* should help readers to appreciate the

⁷ L. Samuelson, *Evolutionary games and equilibrium selection* (Cambridge (MA): MIT Press, 1998).

importance of game theoretical analysis to climate change as well as to other social problems.

Northeastern University

If you need to cite this article, please use the following format:

Smead, Rony and Ronald Sandler, "Game Theory and The Ethics of Global Climate Change," *Philosophy and Public Issues (New Series)*, Vol. 3, No. 1 (2013), 13-23,
edited by S. Maffettone, G. Pellegrino and M. Bocchiola